

1 / 154

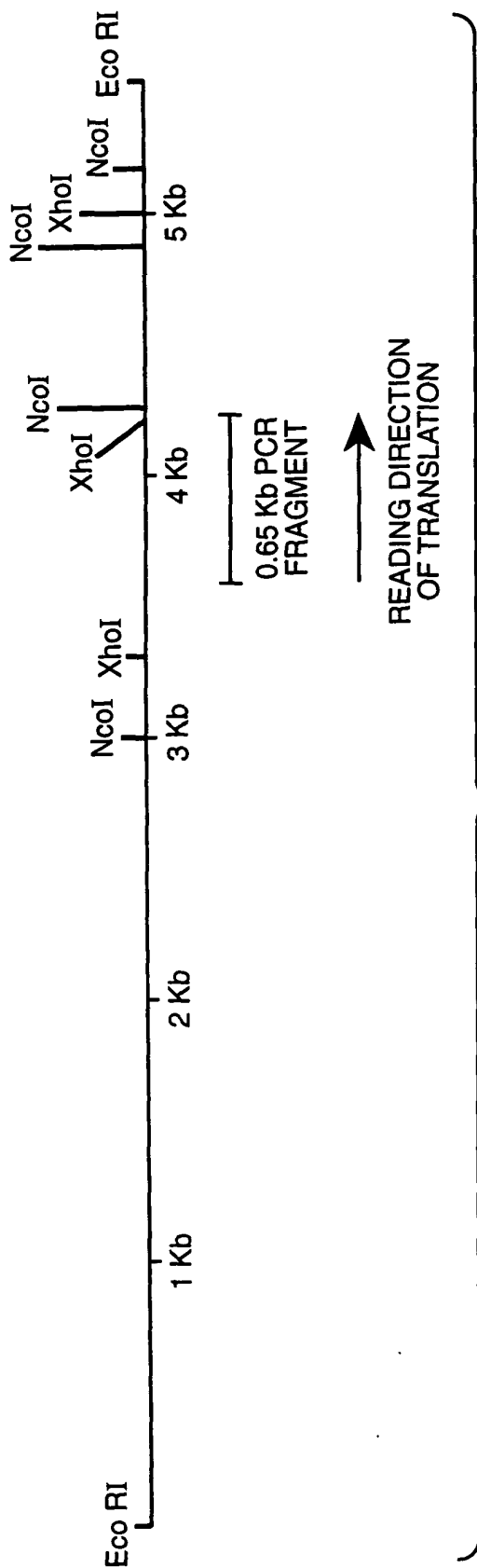


FIG. 1

+

FIG.-2A

3 / 154

Pst I Bpu10 I
ACGTCCTCGCAGTTGTGGTGA CTGCAGGGCACGCCCTTAGCAGCCTCTACGCAAGGCATCTCCGAAGACCTCTACAGCCGTTTAGTCGAAA 630

signal sequence

H V L A V V T A G H A L A A S T O G I S E D L Y S R L V E

Msc I Sal I
TGGCCACTATCTCCCAAGCTGCCCTACGCCGACCTGTGCAACATTCGGTCGACTATTATCAAGGGAGAGAAAAATTTACAATTCTCAAACTG 720

BamH I BsaB I
ACATTACGGATGGATCCTCCGGACGACAGCAGCAAGAAATAATCACCGTCTTCCGGTGGCACTGGTAGTGATACGGAATCTACAACCTCG 810

Eco31 I
ATACTAACTACACCCCTCAGCCCTTTCGACACCCCTACCACAATGCAACGGTTGTGAAGTACACGGTGGATATTATATTGGATGGGTCCTCCG 900

D T N Y T L T P F D T L P Q C N G C E V H G G Y Y I G W V S

FIG._2B

4 / 154

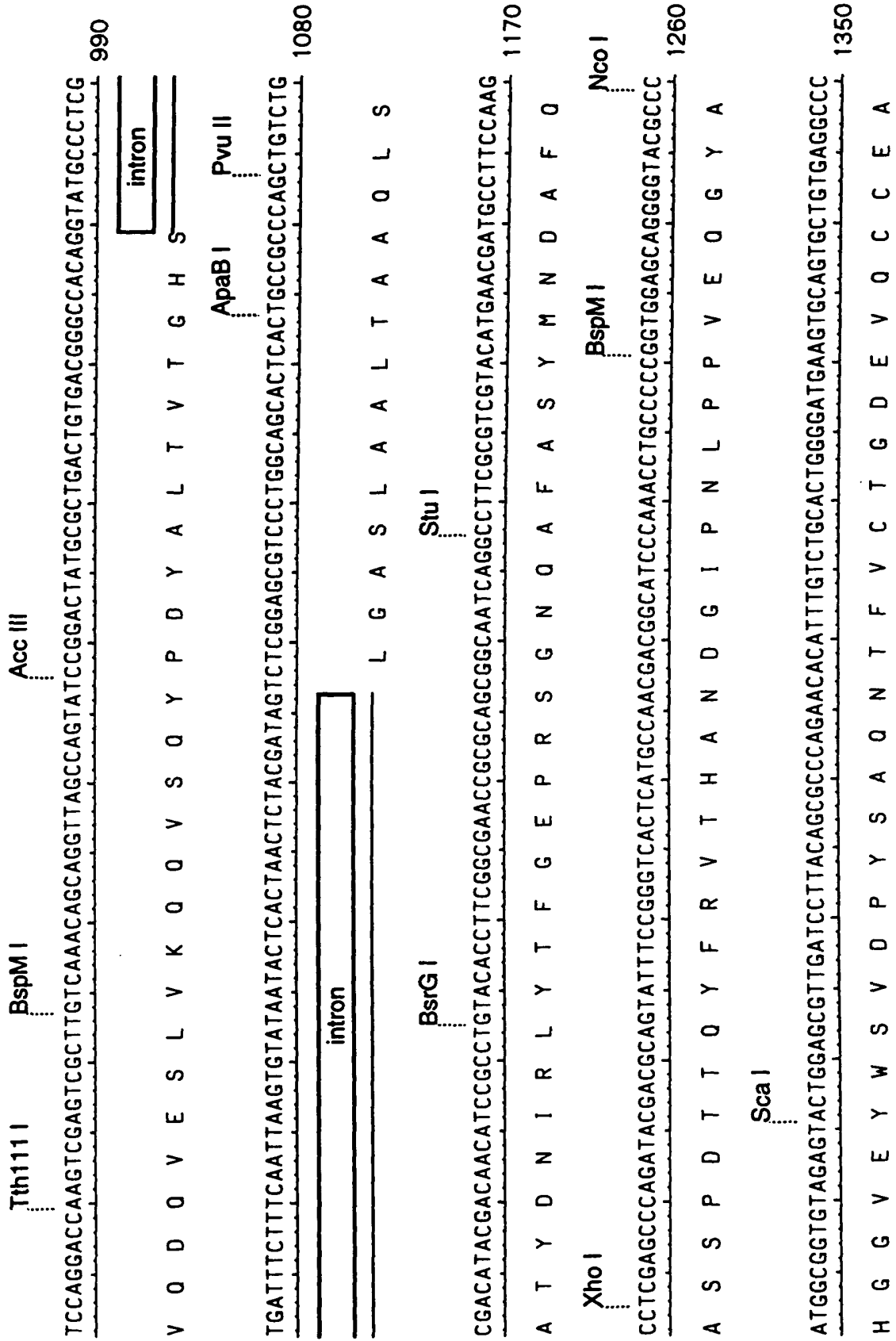


FIG. 2C

5 / 154

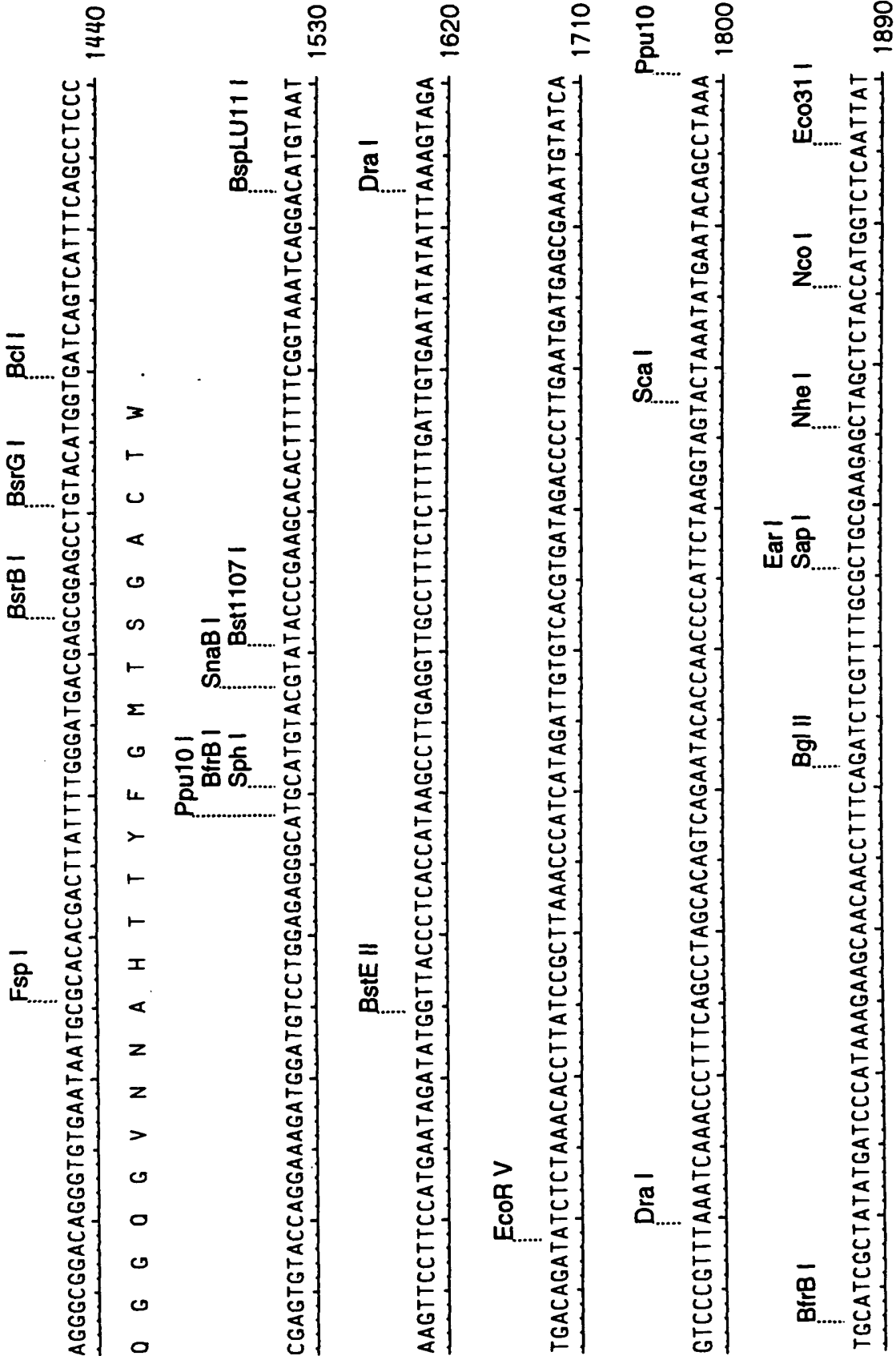


FIG._2D

6 / 154

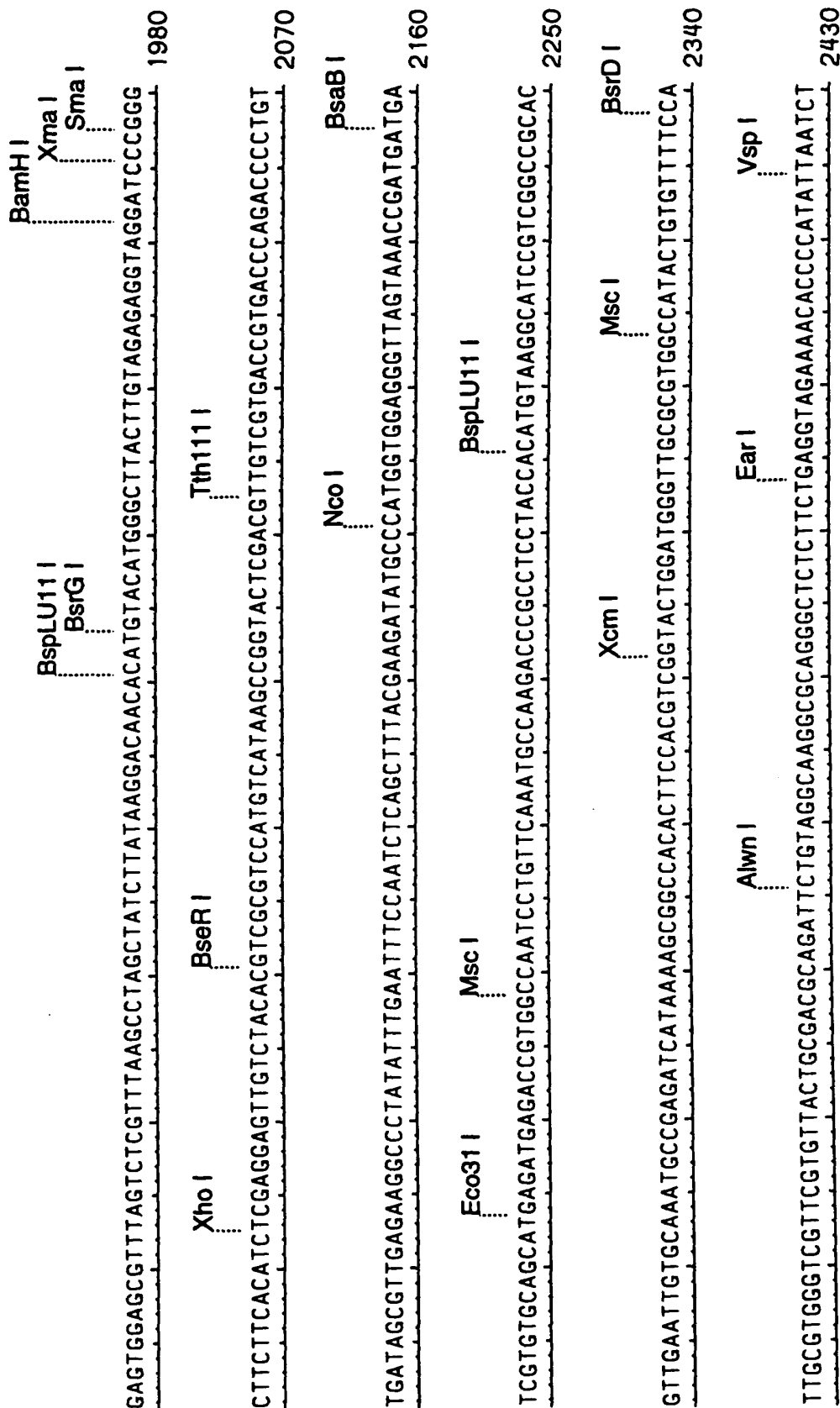


FIG.-2E



7 / 154

CCATGGTGGTGTGATATCGGCAGTAGTCTTTGGCCGAAACGTTGAGGGTTACAGTGATCTGCGTGGACATACATT
CGGGGAATCTACGGCGGAATATCAAGTCTTTCGGAATATCCATATTTGGAAAGGACAGAAAGCTCCGGGGTAGTTT
GATAGTAGCTCCGGGTATTAATCGGGAGCTGACAGGAGTGAGCGTCACTGTAGACCATCTAGTAATGTCAAT
CGCGGCAATTTCCGACATGAAACAAGTTGATTTCCGGGACCCCATTTGTACATCTCTCGGCTACAGCTCGAGATG
TGCTGCGGAGTATCTTAGAGCAGCAGCGTGTGTTTACGACCAAAAGTCAAGGAAATATGAAACGATCG
TCGGATATTTCTTGTATCTTAATAGTCTTCCAGTGGTTTATTAAGAGATAGATCCCTTCACAAACACT
CATCCAACGGACTTCTCATACCACTCATTTGACATAATTTCAAAACAGCTCCAGGCGCATTTAGTTCAACATGAAGC
AATTCCTCCGCAACACGCTCTCGCAGTTGTGGTGACTGCAAGGCAAGCTTACGAGCTCTACGCAAGGCATCT
CCGAAGACCTCTACAGCCGTTTAGTCGAAATGGCCACTATCTCCCAAGCTGCCACGCGGACCTGTGCAACATTC
CGTCGACTATTAACAAGGAGAGAAATTTACAATCTCAAACTGACATTAACGGAATGGAATCTCCGCGAGACA
GCAGCAAGAATAATCACCGTCTTCCGTTGGCACTGGTAGTACGAATCTCAAACTCGATACATACTACACCC
TCACGCTTTCGACACCCCTACCAATGCAACGCTGTGAAAGTACACGCTGGAATATTAATTTGGATGGGTCTCCG
TCCAGGACCAAGTCGAGTCGCTTGTCAAAACAGCAGGTTAGCCAGTATCCGGACTATGCGCTGACTGTGACGGGCC
ACAGGTATGCTTCTGATTTCTTCAATTAAGTGTAATACTCACTAACTCTACGATAGTCTCGGAGCGTCCC
TGGCAGCACTCACTGCGCCAGCTGTCTGCGACATACGACAACTCCGCTGTACACCTTCGGCGAAACCGCGCA
GCGGCAATCAGGCCCTTCGCGTCTGATACATGAACGATGCCCTTCCAAAGCTCCGAGCCAGATACGACGCAGTATTTCC
GGGTCACTCATGCCCCAACGACGGCATCCCCAAACCTGCCCTGGAGCAGGGGTACGCCCATGGCGGTGTAGAGT
ACTGGAGCGTTGATCCTTACAGCGCCCCAGAACACATTTGTCTGCTGCACTGGGGATGAAGTGCAAGTGTGAGGCC
AGGCGGACAGGGTGTGAATAATGCGCACACGACTTATTTTGGGATGACGAGCGGAGCTGTACATGTTGATCAG
TCATTTACGCTCCCGAGTGATCCAGGAAGATGGATGTCTGGAGAGGGCATGCAATGATACGTATACCCGAAAGC
ACACTTTTTCGGTAAATCAGGACATGTAATAAGTTCCCTTCCATGAATAGATATGGTTACCTCACCATAAGCCTT
GAGGTGCTTCTTCTTGTGATTTGTGAATATATATTTAAAGTAGATGACAGATATCTCTAAACACCTTATCCGCT
TAAACCCATCATAGATTGTGTCAAGTATAGACCCCTTGAATGATGAGCGAAATGTATCAGTCCCGTTTAAATCA
AACCCTTTCAGCTAGCACAGTCAGAAATACACCAACCCCATTTCTAAGGTAGTACTAAATATGAATACAGCCTAAA
TGCAATCGCTATATGATCCCATAAAGAGCAACAACTTTCAGATCTCGTTTTCGCTGCGAAAGAGCTAGCTCTAC
CATGGTCTCAATTATGAGTGGAGCGTTTAGTCTCGTTTAAAGCTTAGCTATCTTATAAGGACAAACATGTACATG
GGCTACTGTAGAGGTAGGATCCCGGGCTTCTTCACATCTCGAGGAGTTGTCTACACGTCGCGTCCATGTCA
TAAGCCGGTACTCGACGTTGTCTGTGACCGTGACCCAGACCCCTGTTGATAGCGTTGAGAGGCCCTATATTGAA
TTTCCAAATCTCAGCTTACGAAAGATATGCCCATGTTGAGGGTTAGTAAACCCGATGATGATGTCGTCAGCATGA
GATGAGACCGTGGCCAAATCCCTGTTCAAAATGCCAAAGACCCGCTCCCTACCACATGTAAAGCATCCGTCGGCCGAC
GTTGAATTGTGCAATGCCGAGATCATAAAGCGGCCACACTTCCACGTCGGTACTGGATGGGTGCGCGGTGGCC
ATACGTGTTTTCATGCGTGGGTGTTCTGCGACGCAATTTCTGTAGGCAAGGCGCAGGCGCTCTCT
TCTGAGGTAGAAACACCCCATATTAATCTGAATC

FIG._3



40-mer

•

complement, FAE-I5
FAE-I3

FIG. 5

9 / 154

Vector Construction

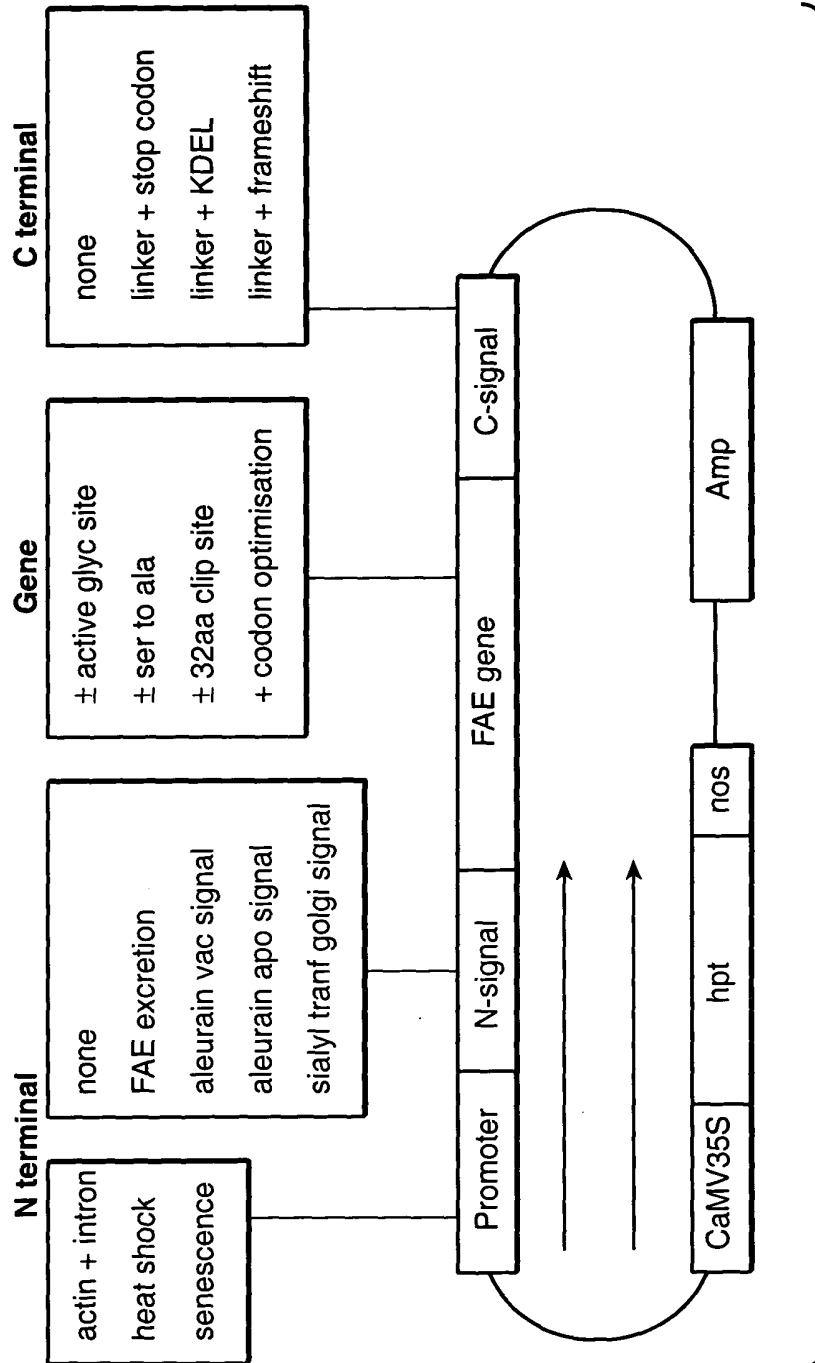


FIG._6

10 / 154

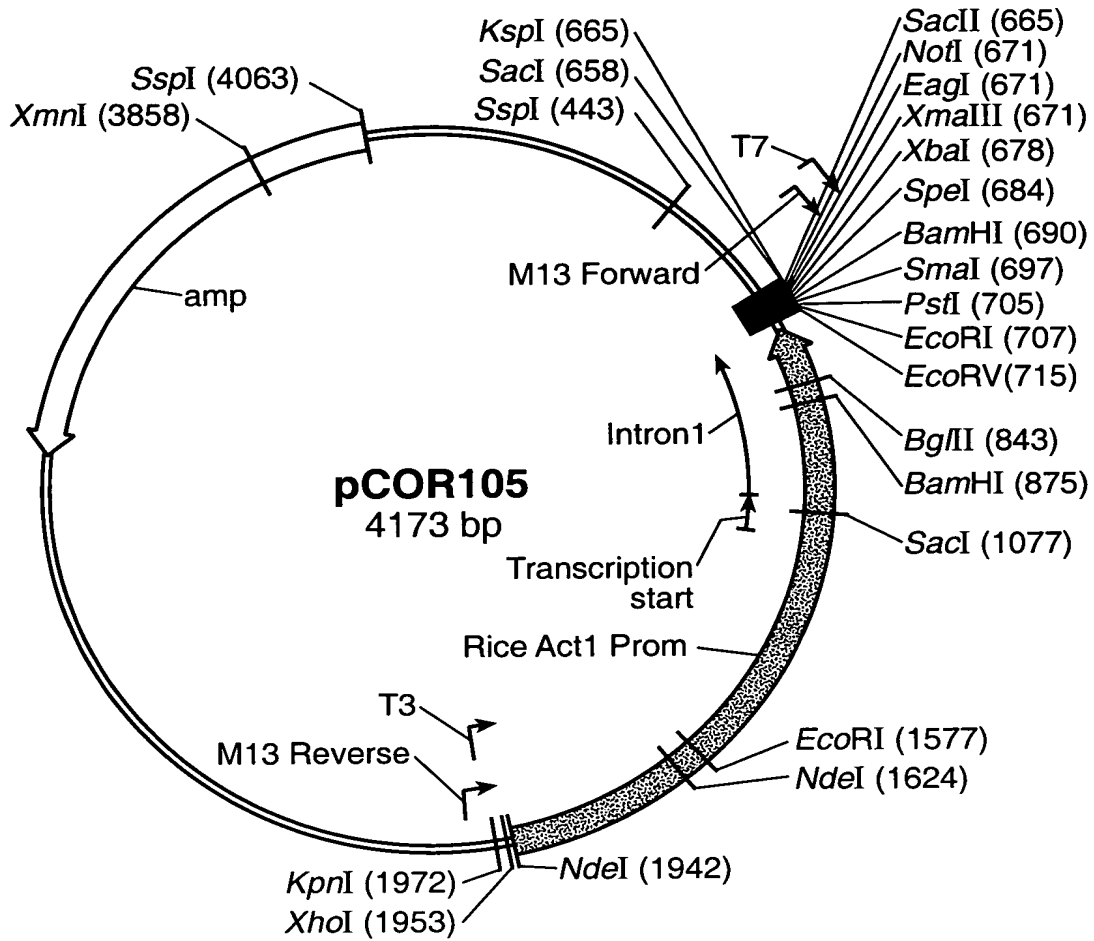


FIG._7

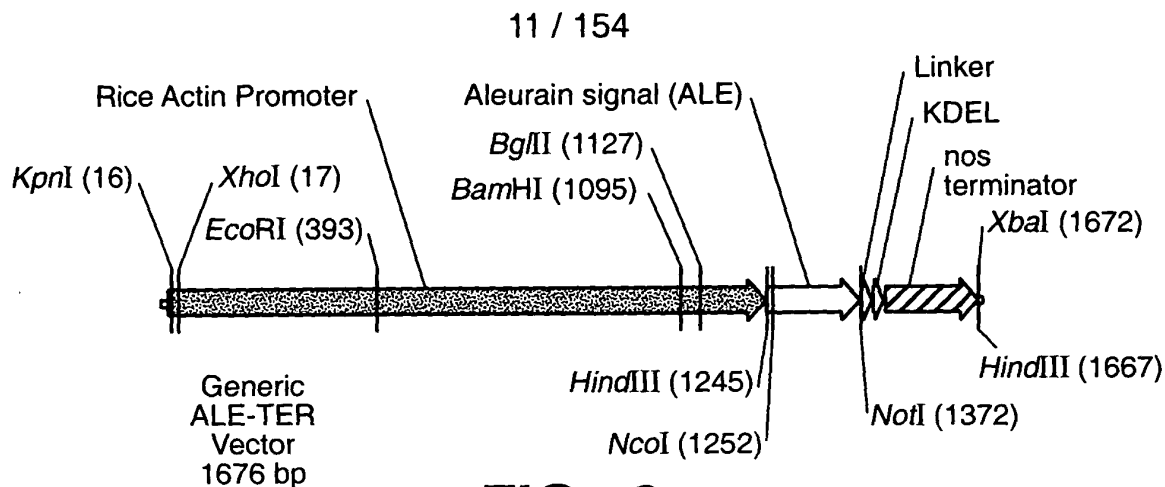


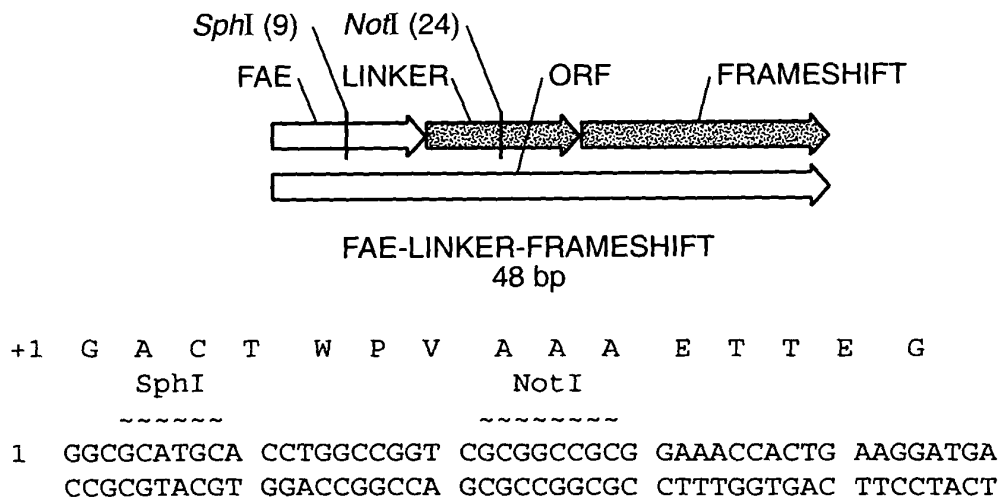
FIG._8

KDEL-COOH ER retention sequence

NotI
 ~~~~~  
 A A A K P L K D E L \*  
 1 GCGGCCGCGA AACCACTGAA GGATGAGCTG TAA

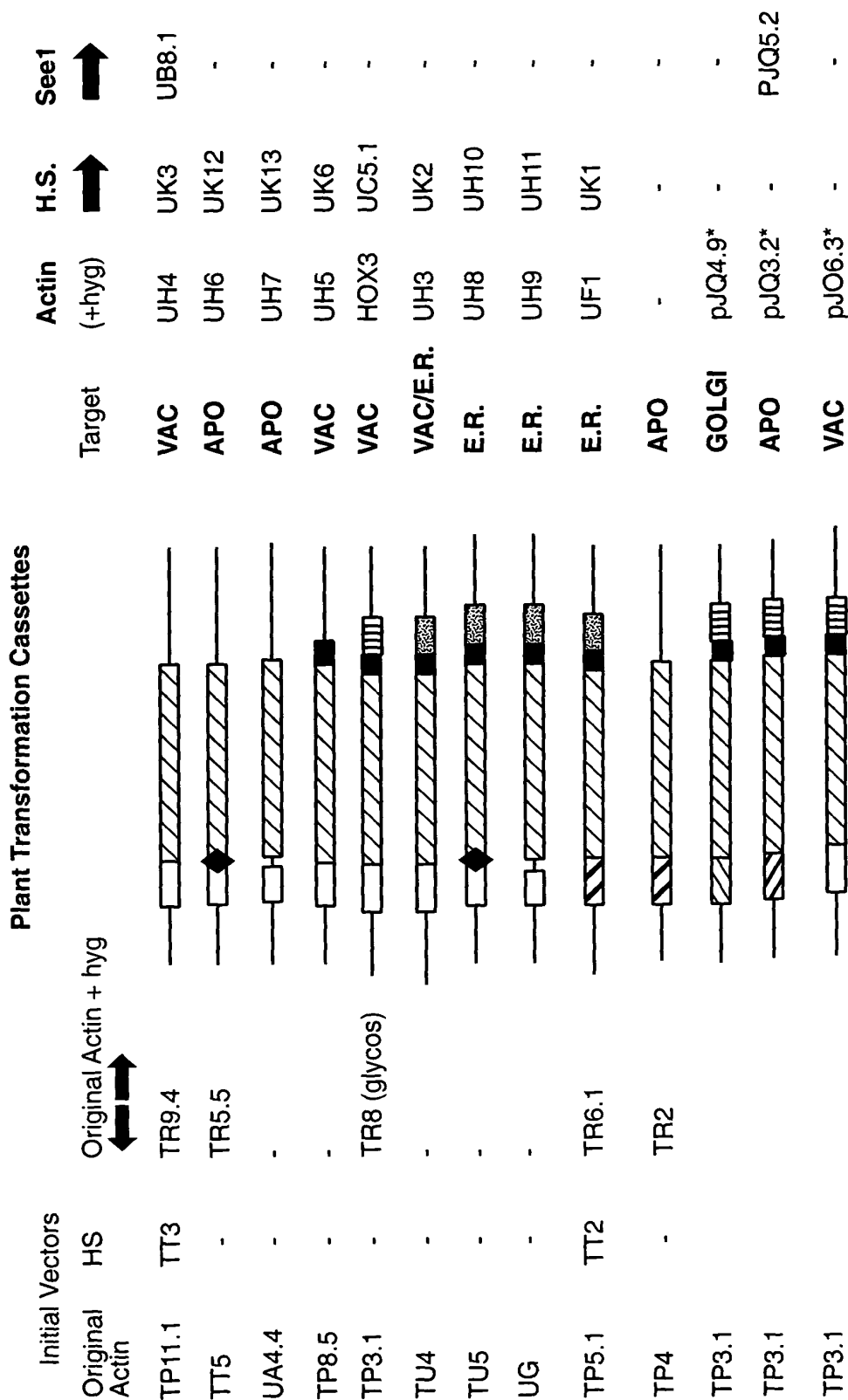
**FIG.\_9**

**FAE-LINKER-FRAMESHIFT Structure and Sequence**

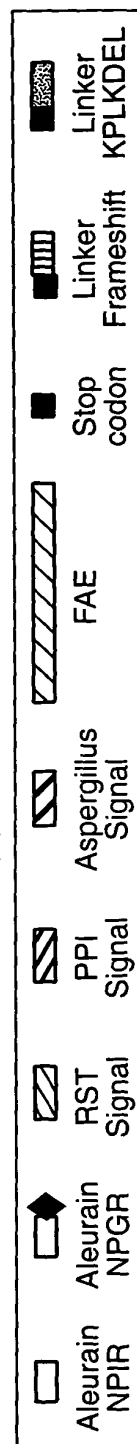


**FIG.\_10**

12 / 154



\* - Modified Actin Promoter (Kpn1-EcoR1 Deletion and Restored NCO Site)



**FIG.-11**

13 / 154

## Vectors

### Original Actin promoter in pCOR105

|       | Target | Signal sequences                      | Vectors                      |
|-------|--------|---------------------------------------|------------------------------|
| (i)   | APO    | - aleurain-NPGR-FAE                   | pUH6, pTT5, TT5.5, pTT5.1    |
|       |        | - aleurain-delNPIR -FAE               | pUH7, pUA4.4,                |
| (ii)  | ER     | - aleurain-NPGR-FAE-linker-KDEL       | pTU5, pUH8,                  |
|       |        | - aleurain-delNPIR-FAE-linker-KDEL    | pUG4, pUH9,                  |
| (iii) | VAC    | - aleurain-NPIR-FAE                   | pTP11.1, pTR9.4, pUH4, pUK3, |
| (iv)  | ER/VAC | - aleurain-NPIR-FAE-linker-KDEL       | pTU4, pUH3,                  |
| (v)   | VAC    | - aleurain-NPIR-FAE-linker-frameshift | pUA1K3, pTP3.1, pUC5.11      |
| (vi)  | VAC    | - aleurain-NPIR-FAE-linker-stop       | pTP8.5, pUH5                 |
| (vii) | ER     | - Aspergillus signal -FAE-KDEL        | pTP5.1, pTP6.1, pUF1,        |

### Modified actin promoter (Kpn1-EcoR1 deletion and restored NCO site)

|       |       |                                       |        |
|-------|-------|---------------------------------------|--------|
| (i)   | VAC   | - aleurain-NPIR-FAE-linker-frameshift | pJ06.3 |
| (ii)  | GOLGI | - RST-FAE-linker-frameshift           | pJQ3.2 |
| (iii) | APO   | - PPI-FAE-linker-frameshift           | pJQ4.9 |

### Heat-shock promoter

|       |        |                                       |                  |
|-------|--------|---------------------------------------|------------------|
| (i)   | APO    | - aleurain-NPGR-FAE                   | pUH12            |
|       |        | - aleurain-delNPIR-FAE                | pUH13            |
|       |        | - Aspergillus signal-FAE              | pTP4a2, pTR2.22, |
| (ii)  | ER     | - aleurain-NPGR-FAE-linker-KDEL       | pUH10            |
|       |        | - aleurain-delNPIR-FAE-linker-KDEL    | pUH11            |
| (iii) | VAC    | - aleurain-NPIR-FAE                   | pUK3, pTT3       |
| (iv)  | ER/VAC | - aleurain-NPIR-FAE-linker-KDEL       | pUK2             |
| (v)   | VAC    | - aleurain-NPIR-FAE-linker-frameshift | pUC5.11, pHOX3   |
| (vi)  | VAC    | - aleurain-NPIR-FAE-linker-stop       | pUK6             |
| (vii) | ER     | - Aspergillus signal -FAE-KDEL        | pUK1, pTT2       |

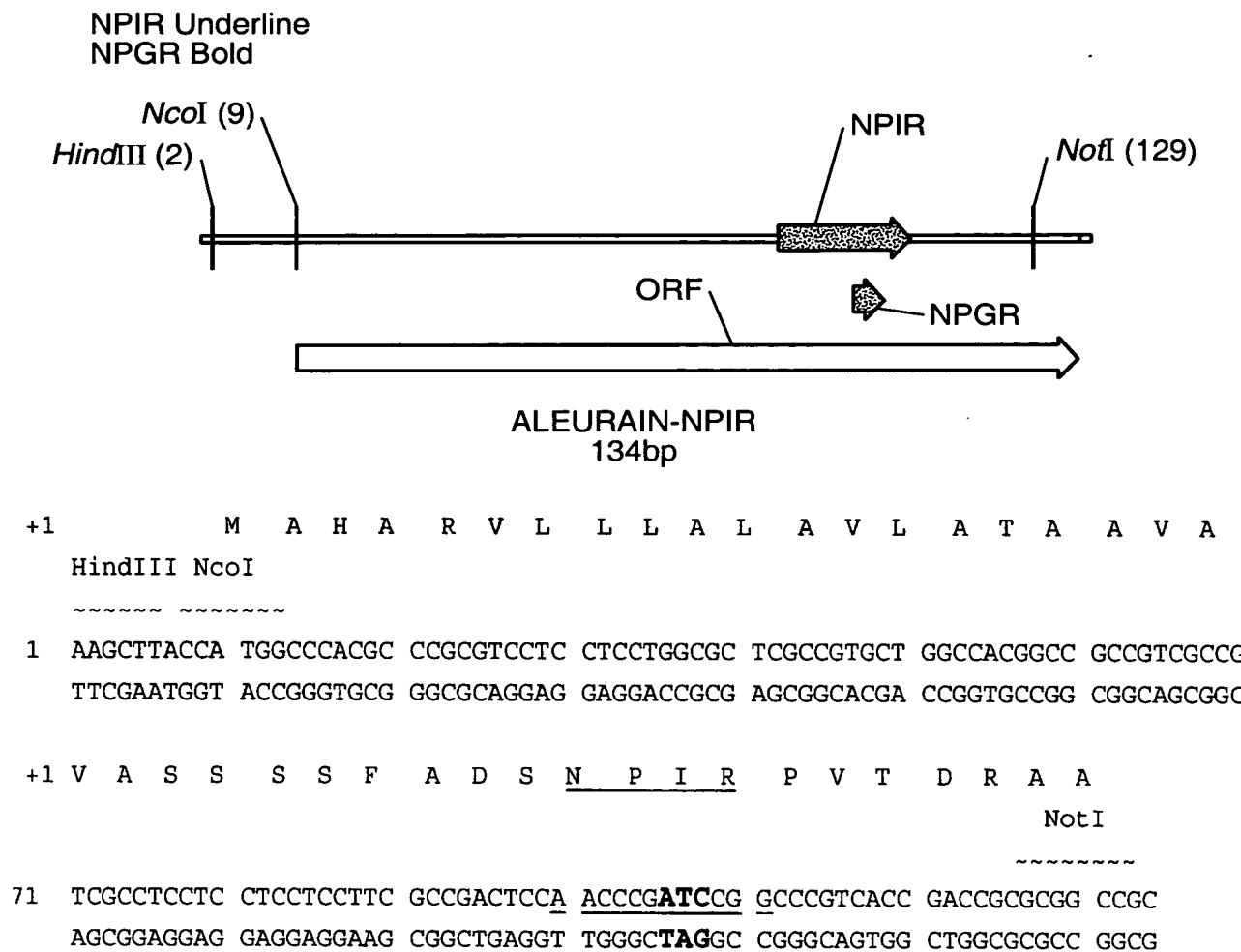
### Senescence promoter

|      |     |                                  |        |
|------|-----|----------------------------------|--------|
| (i)  | APO | - See1-PPI-FAE-linker-frameshift | pJQ5.2 |
| (ii) | VAC | - See1-aleurain-deleted NPIR-FAE | pUB8.1 |

**FIG.\_12**

14 / 154

# **ALEURAIN-NPIR (Vacuolar) and NPGR (Apoplast) Structure and Sequence**



**FIG. 13**

# RAT SIALYL TRANSFERASE Golgi signal sequence

HindIII  
 ~~~~~  
 1 AAGCTTACCA TGATCCACAC CAACCTCAAA AAGAAGTTCT CCTCTTTCAT CCTCGTCTTC CTCCTCTTCG
 M I H T N L K K K F S L F I L V F L L F A
 71 CCGTGATCTG CGGTGTGGAAG AAGGGCTCCG ACTACGAGGC CCTCACCCCTC CAAGCCAAGG AGTTCCAAAT
 . V I C V W K K G S D Y E A L T L Q A K E F Q M
 NotI
 ~~~~~  
 . A A  
 141 GGCGGCCCGC

# FIG.-14

# POTATO PROTEASE INHIBITOR II Apoplast signal sequence

HindIII  
~~~~~  
1 AAGCTTACMA TGGMCGTGCA CAAGGAGGTS AACTTCGTSG CCTACCTCCT GATCGTSCTC
GGCCTCCTCT

NCOI
~~~~~  
71 TGCTCGTSTC CGCCATGGAG CACGTGGAG CCAAGGCCCTG CACCCKCGAG TCGGGCAACC  
TCGGCTTCGG

NotI  
~~~~~  
141 CATCTGCCCG GCGGCGGCC

FIG 15

FIG. 15

16 / 154

Targ ting Expression of gfp to Differ nt Cell Compartm nts

Actin Promoter Targeting Vectors

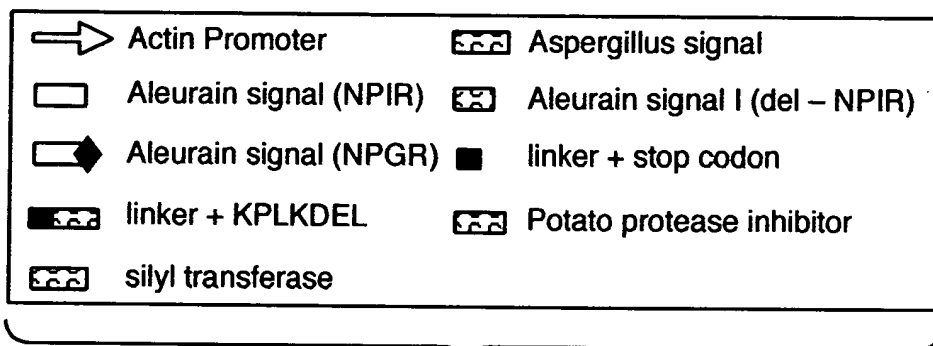
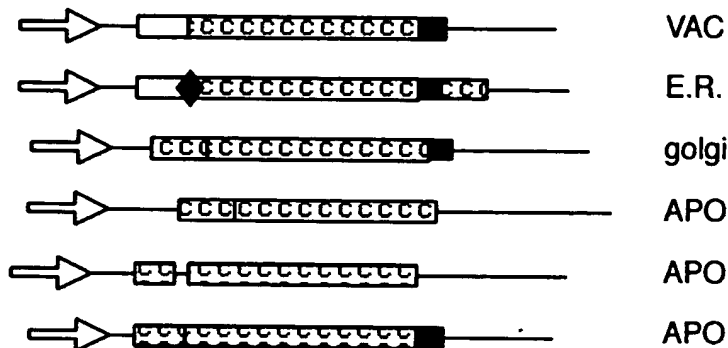


FIG._16A

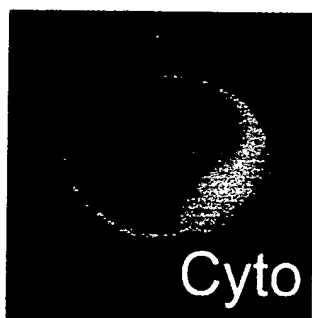


FIG._16B

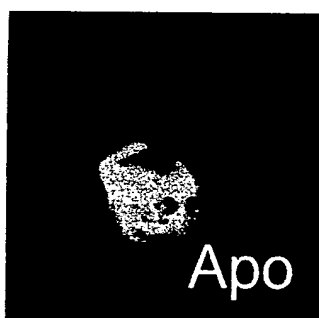


FIG._16C

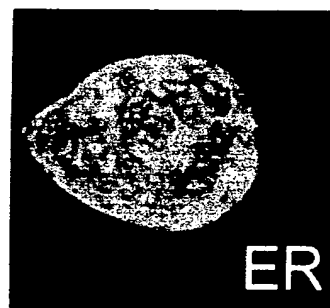


FIG._16D

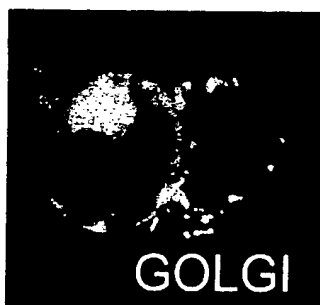


FIG._16E

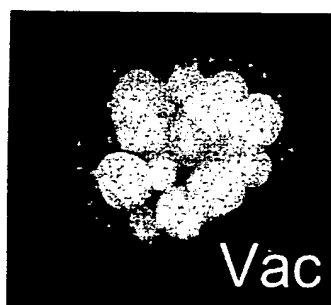


FIG._16F

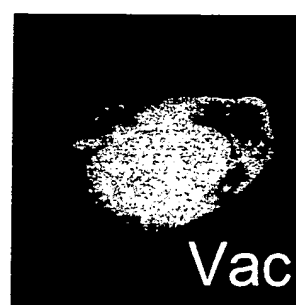


FIG._16G

17 / 154

**FAE Activity in Transgenic *Festuca arundinacea* Leaves
of Different Ages Under ER and APO Targeting Sequence**

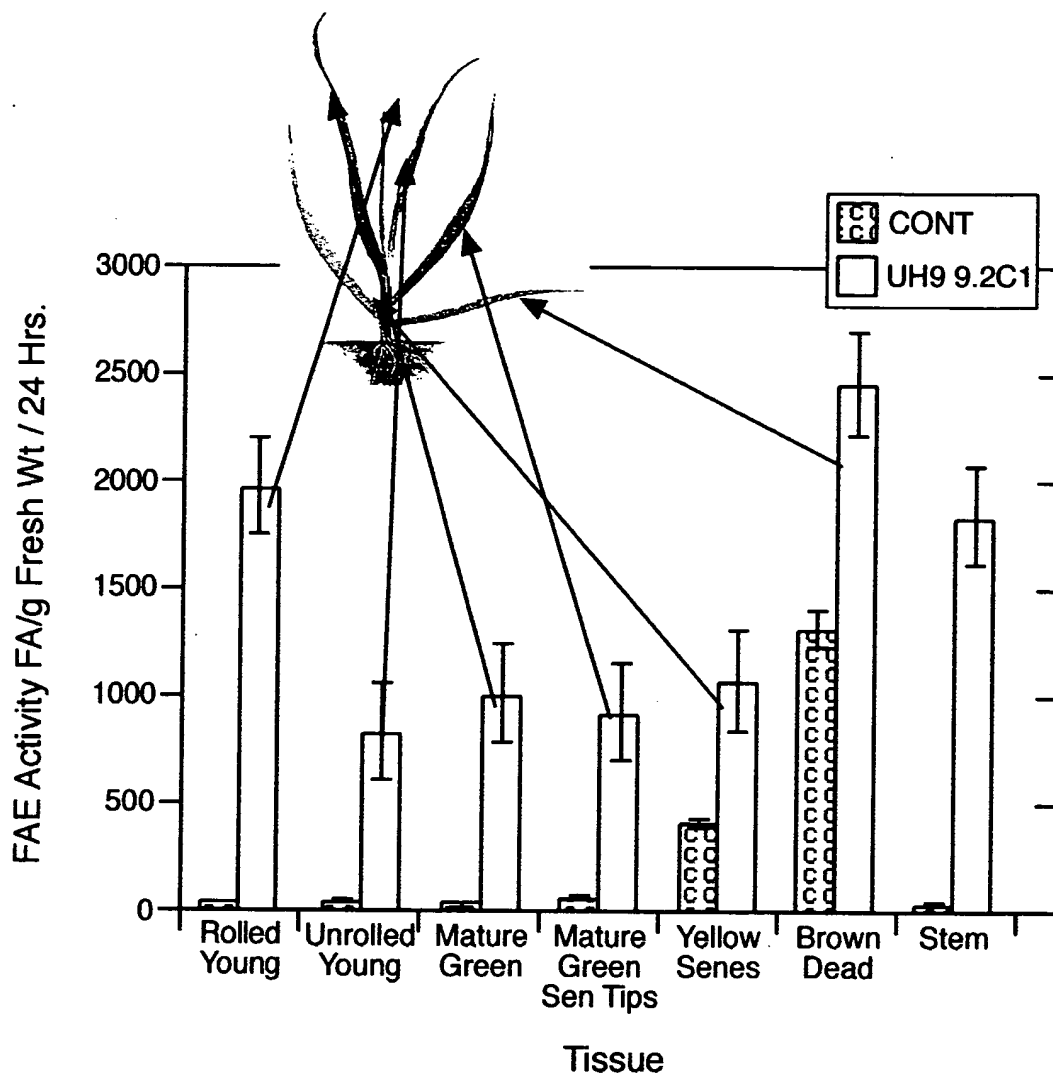


FIG. 17A

18 / 154

FAE Activity in Transgenic *F. stuebelii* Leaves of Different Ages Under ER and APO Targeting Sequence

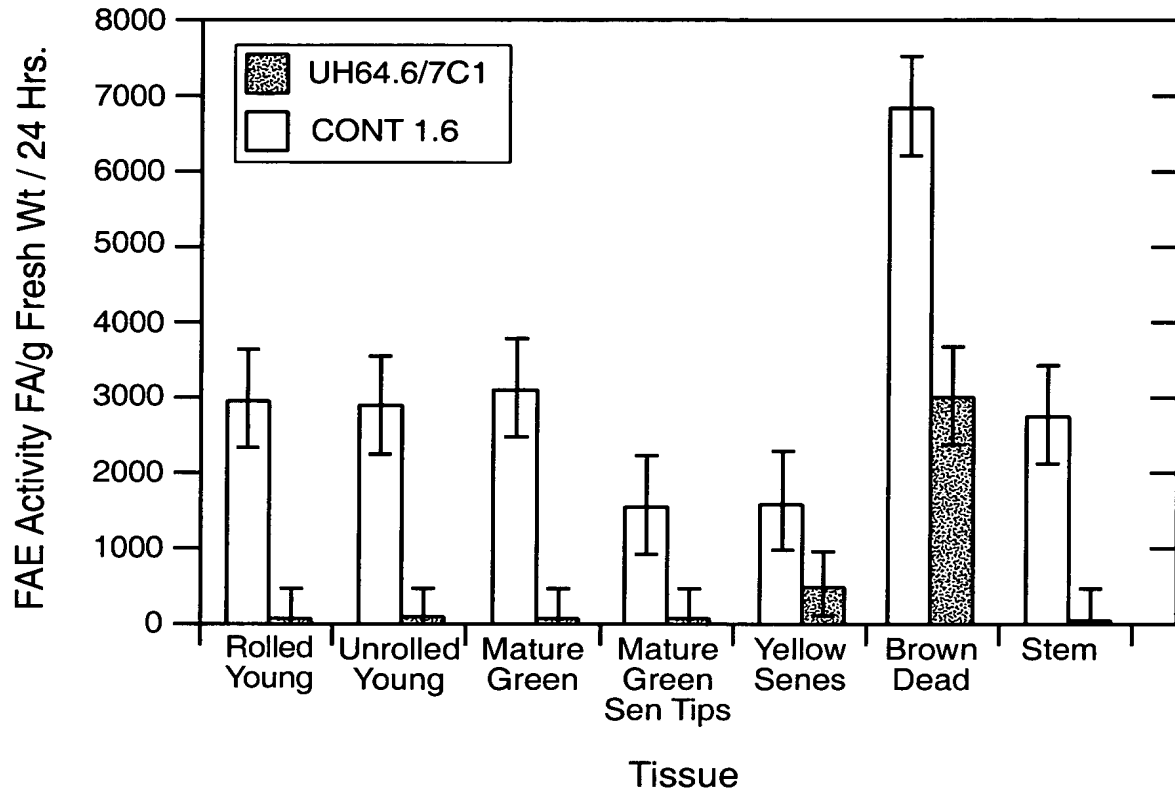


FIG. 17B

19 / 154

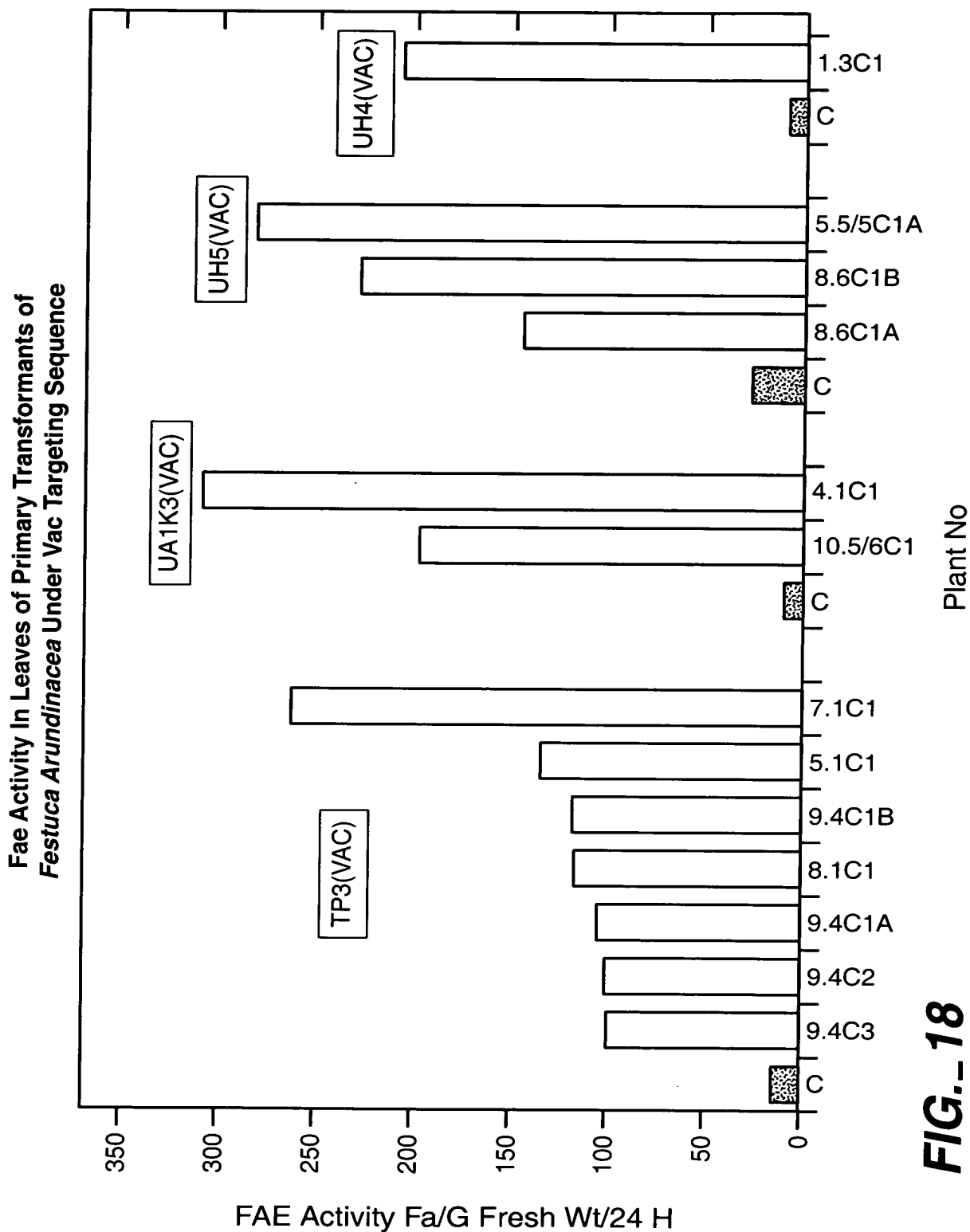
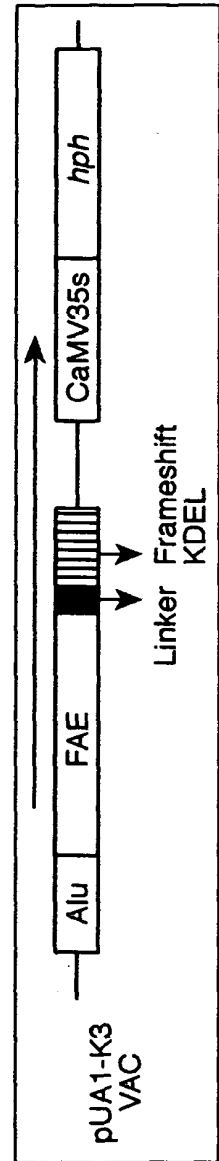
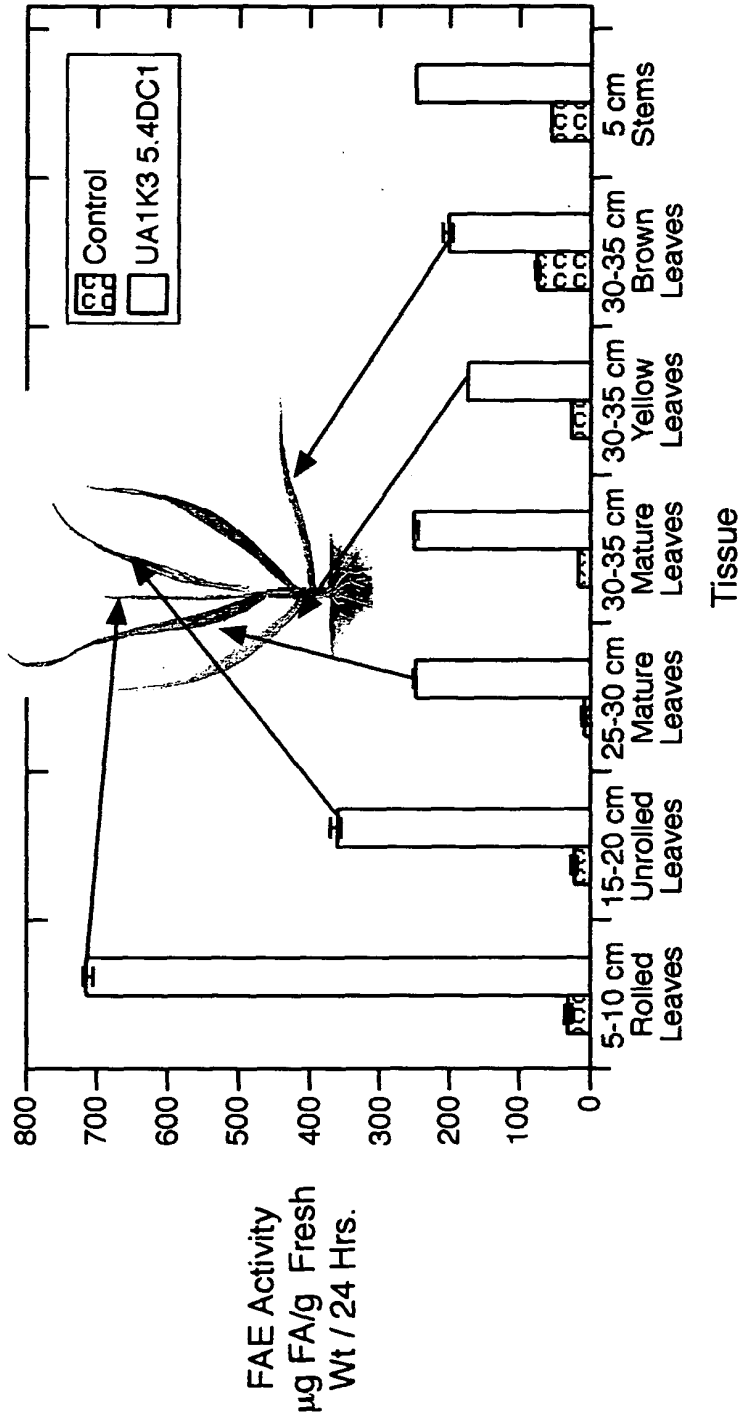


FIG. 19

FAE Activity in *Lolium multiflorum* Leaves of Different Ages



21 / 154

FAE Activity in Leaves of Primary Transformants of *Lolium multiflorum* Under VAC APO and ER Targeting Sequence

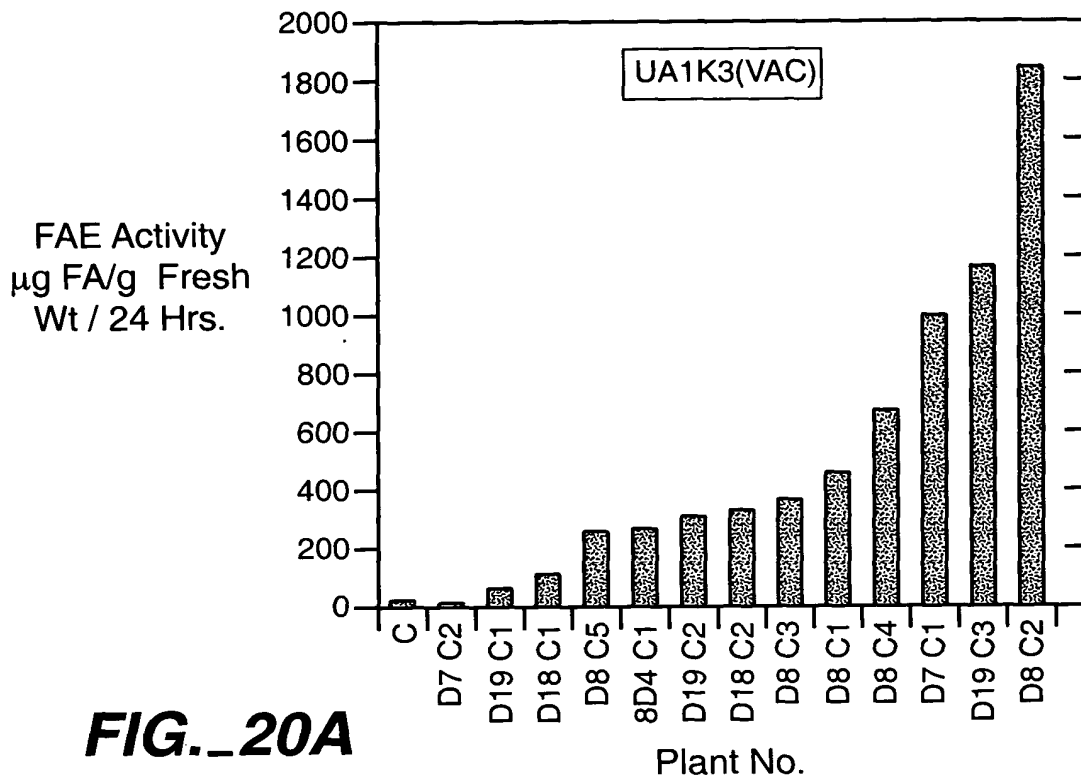


FIG._20A

FAE Activity in Leaves of Primary Transformants of *Lolium multiflorum* Under VAC APO and ER Targeting Sequence

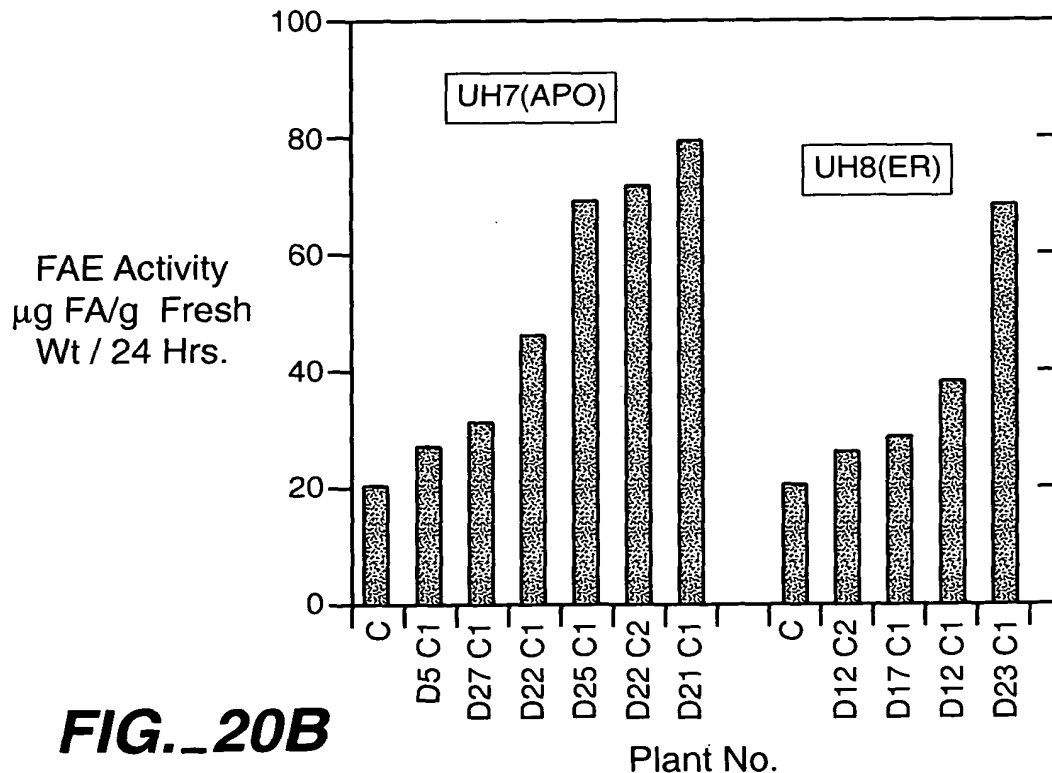


FIG._20B

22 / 154

Levels of Esterified Monomeric and Dimeric Hydroxycinnamicacids in *Festuca Arundinacea* Plants Expressing FAE Under VAC Targeting Sequence

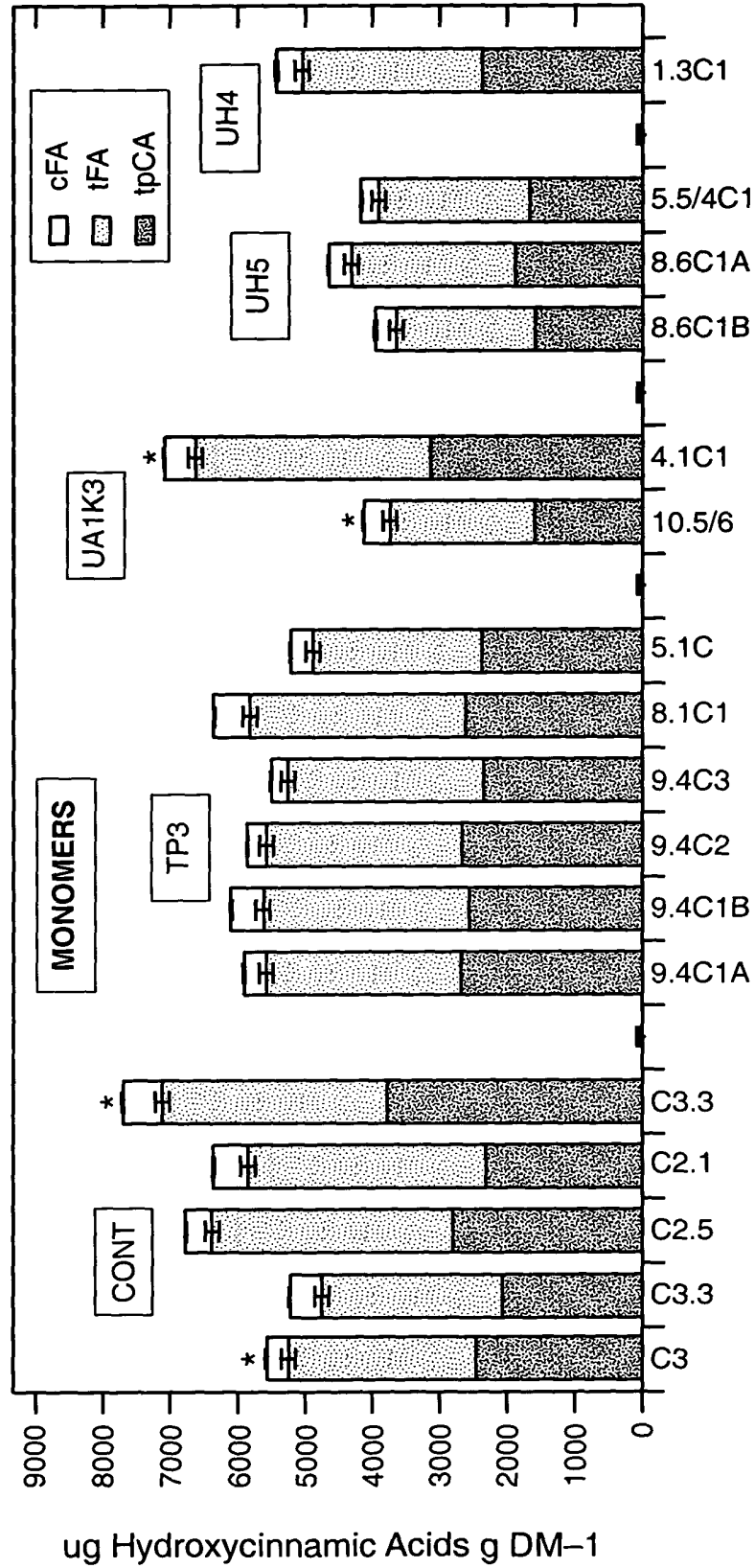


FIG..21A

23 / 154

Levels of Esterified Monomeric and Dimeric Hydroxycinnamicacids in
Festuca Arundinacea Plants Expressing FAE Under VAC Targeting Sequence

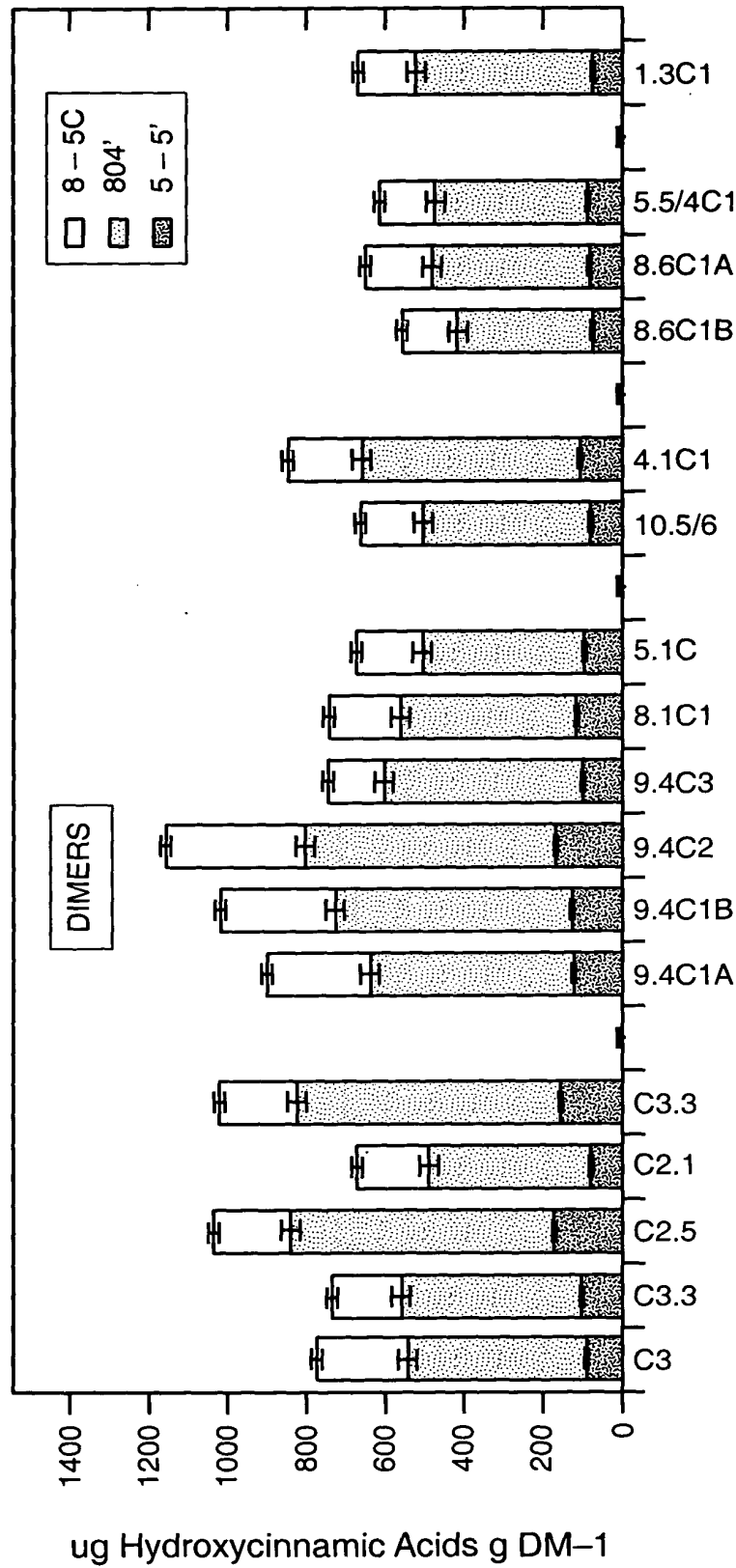


FIG..21B

24 / 154

Levels of Esterified Monomeric and Dimeric Hydroxycinnamic Acids in Leaves of *F. a.* Expressing FAE Under ER and APO Targeting Sequence

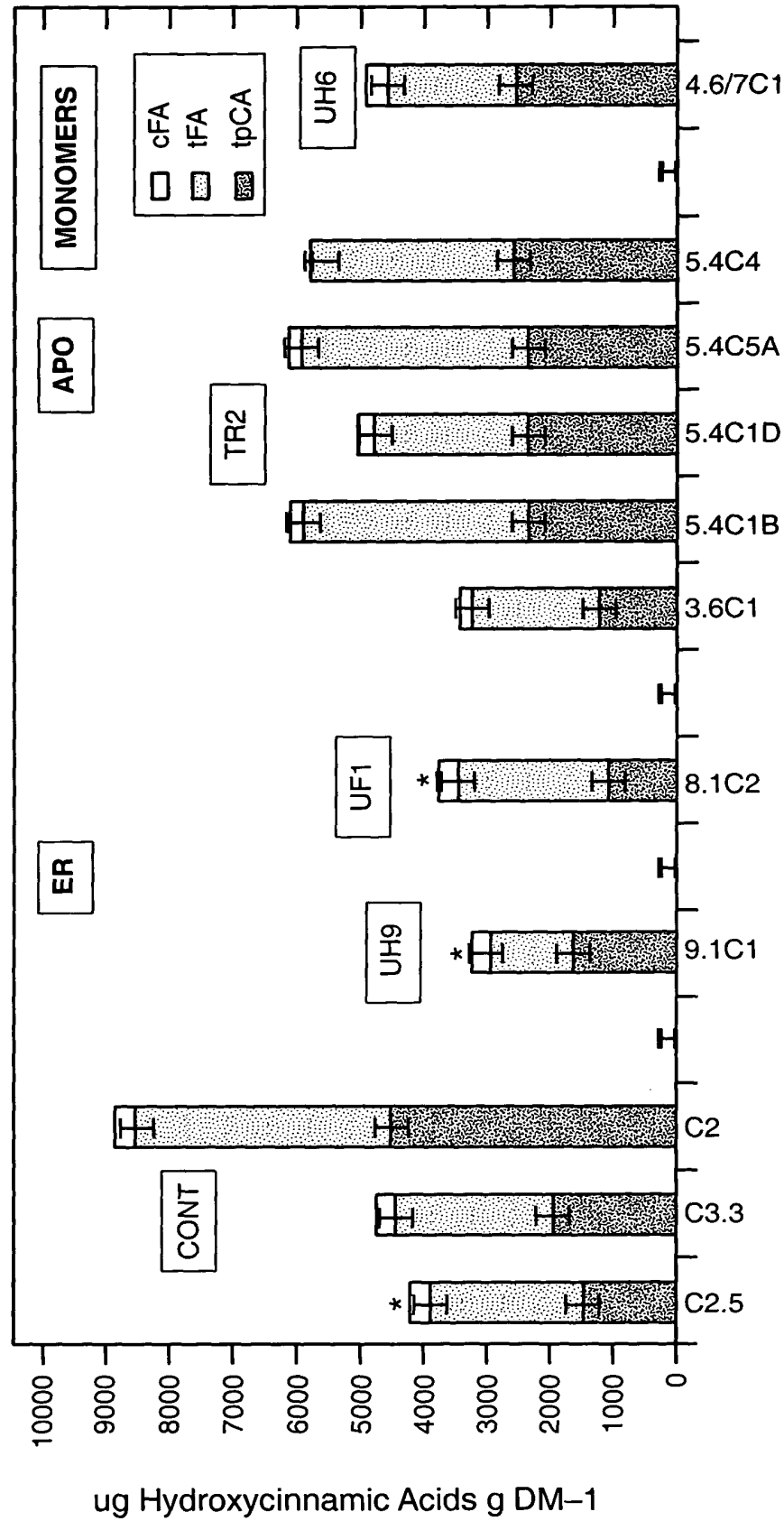


FIG._22A

25 / 154

Levels of Esterified Monomeric and Dimeric Hydroxycinnamic Acids in Leaves of *F. a.* Expressing FAE Under ER and APO Targeting Sequence

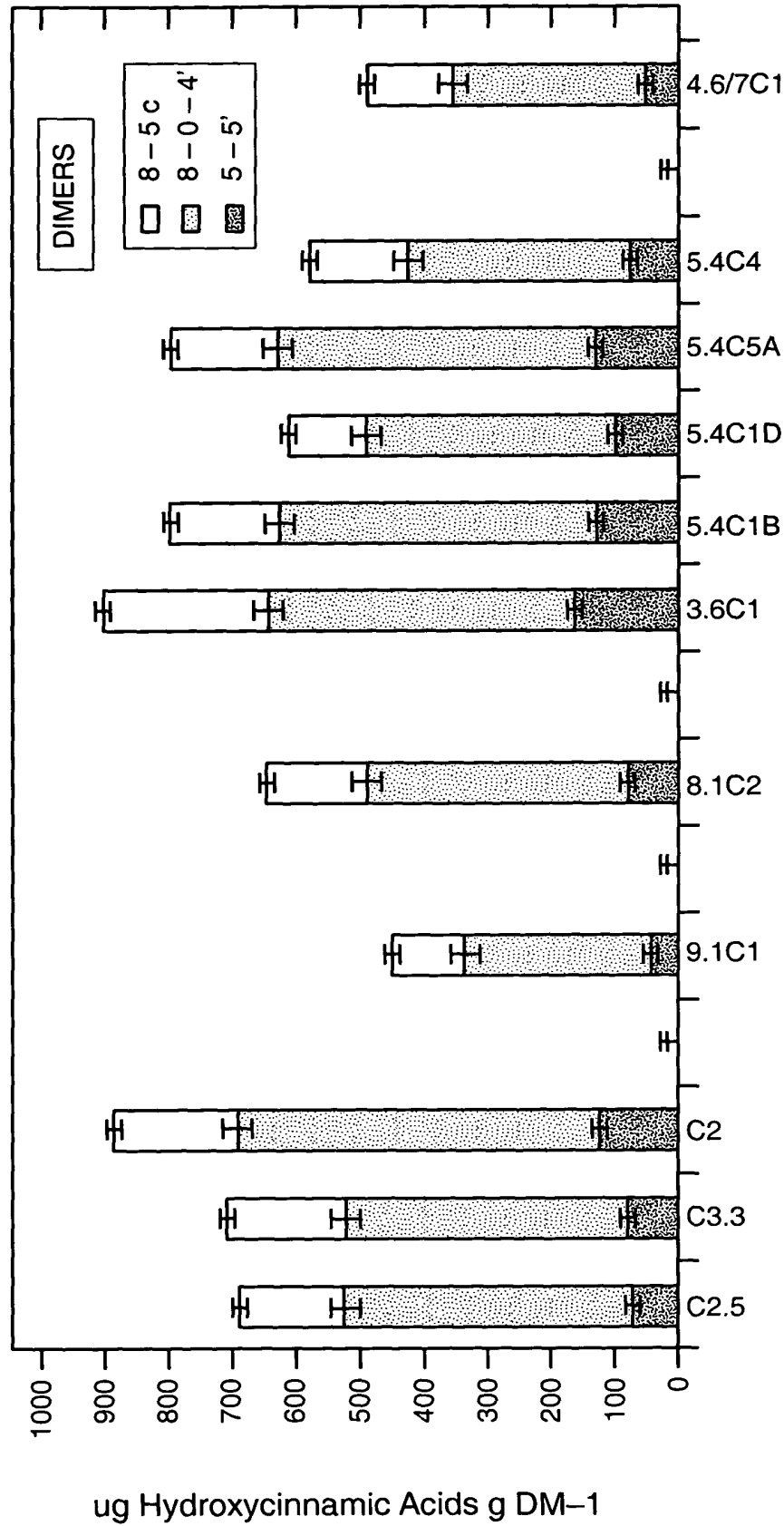


FIG. 22B

26 / 154

In Vitro Dry Matter Digestibility of Leaf Tissue of Mature *Festuca arundinacea* Plants Expressing FAE Under an Actin Promoter

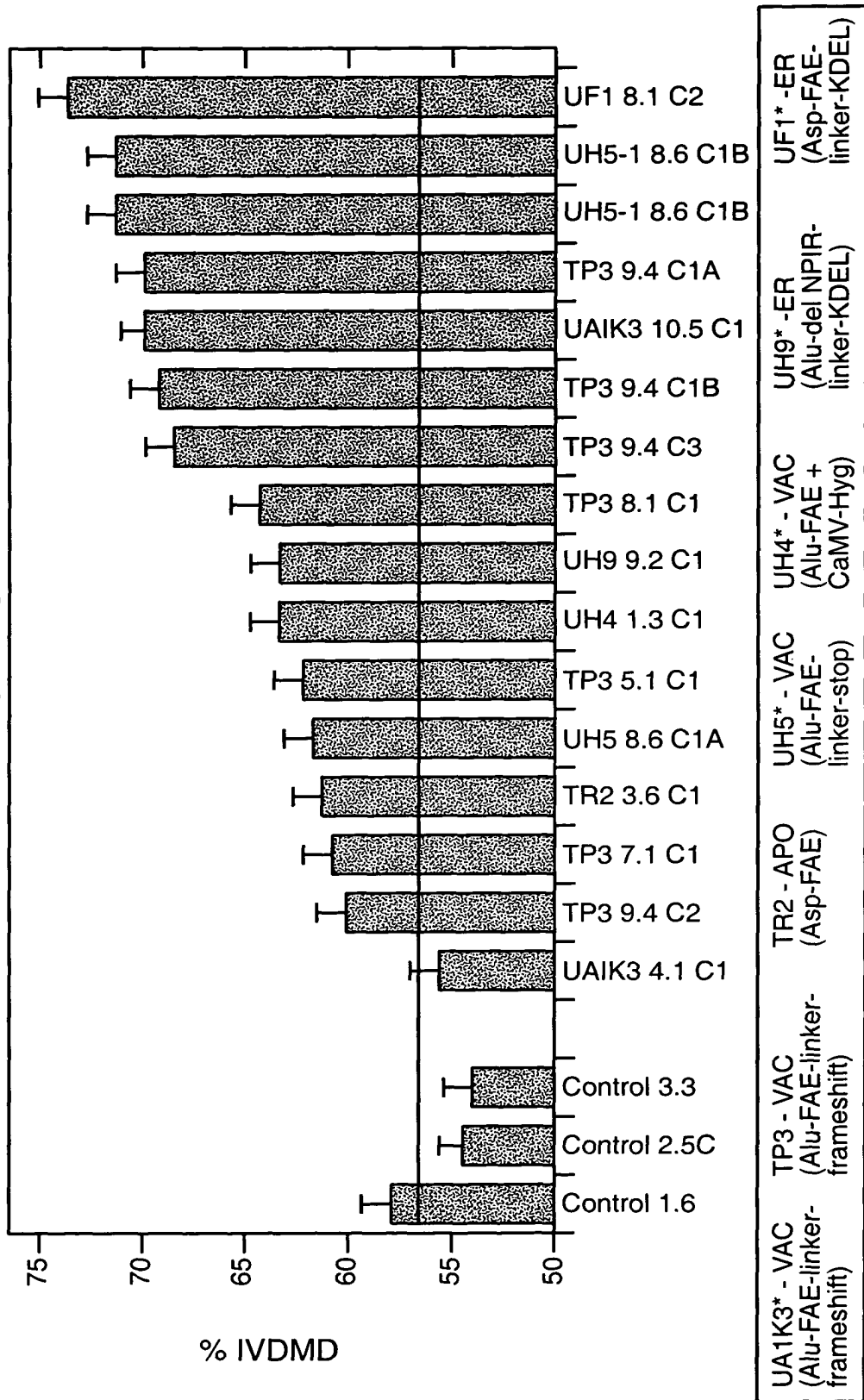


FIG. 23

27 / 154

In Vitro Dry Matter Digestibility of Leaf Tissue of Mature
Lolium multiflorum Plants Expressing FAE Under an Actin Promoter

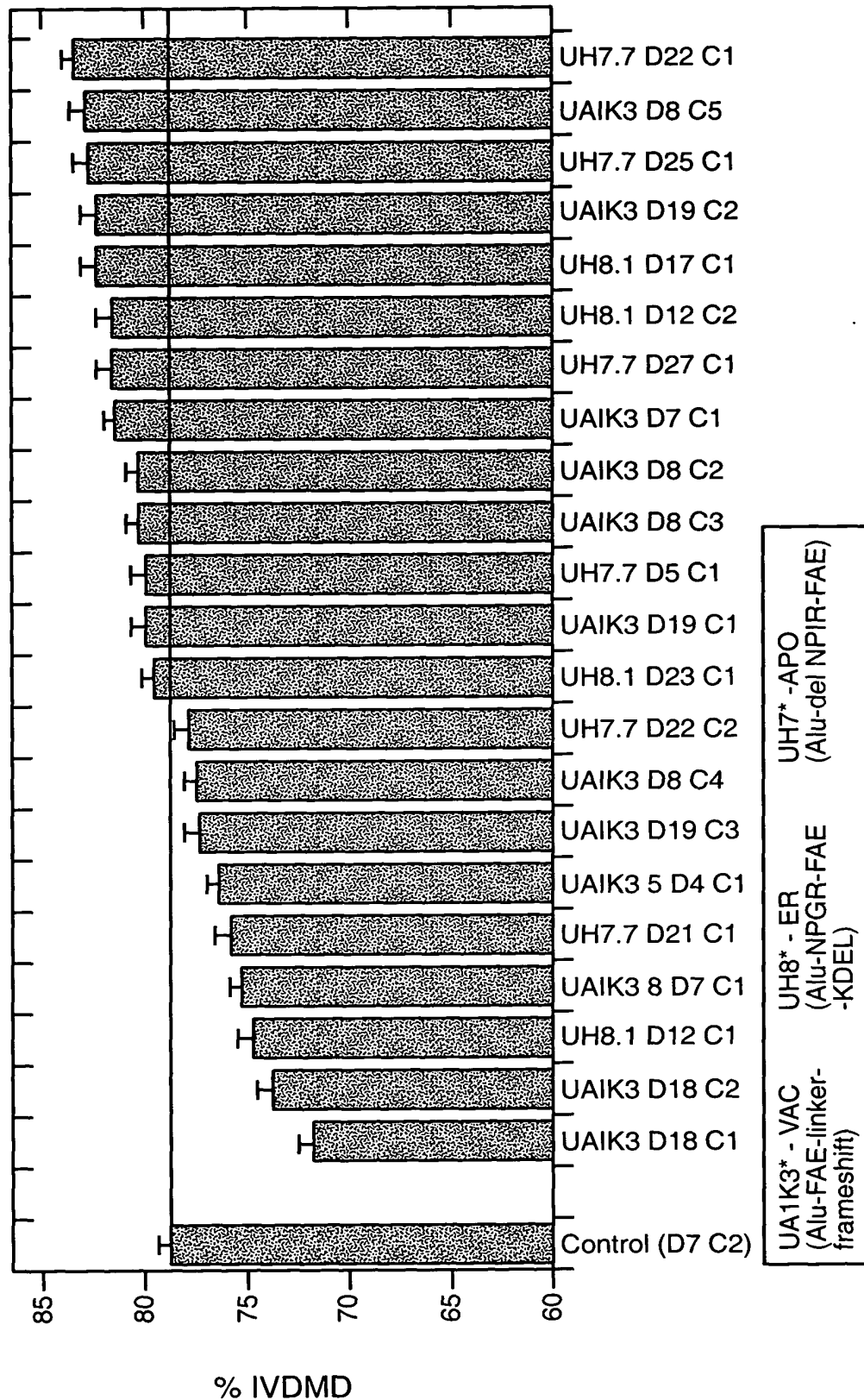


FIG. 24

28 / 154

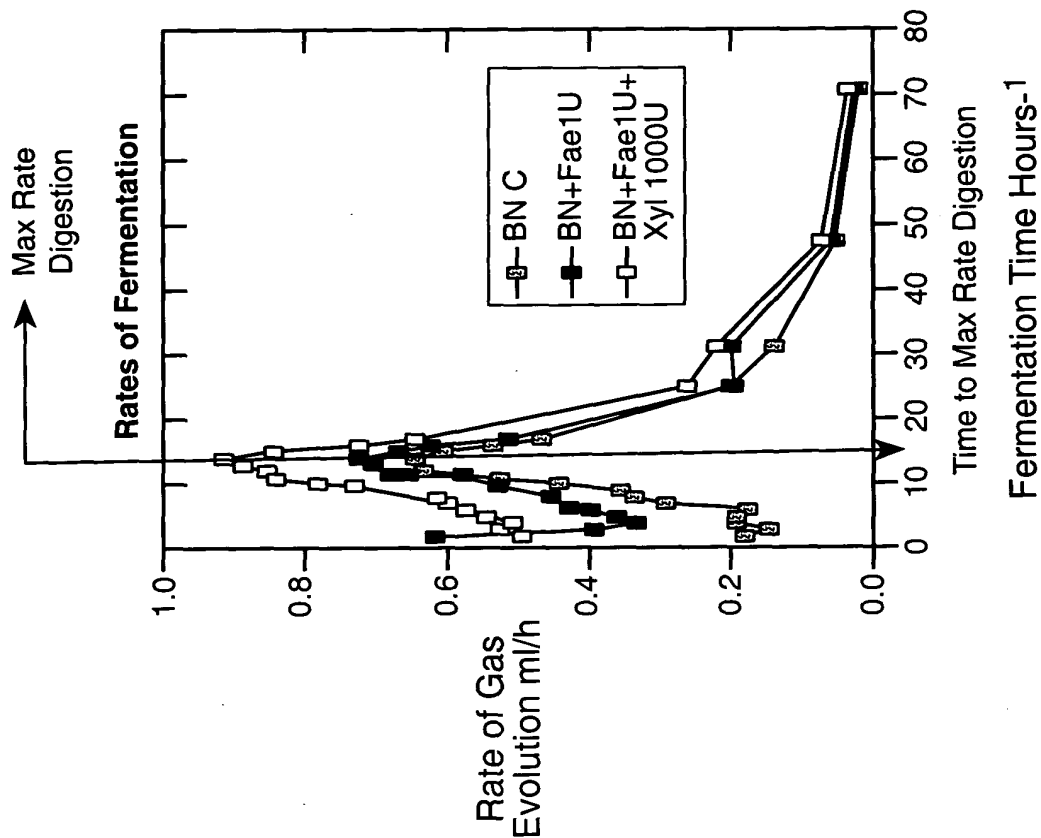


FIG..25B

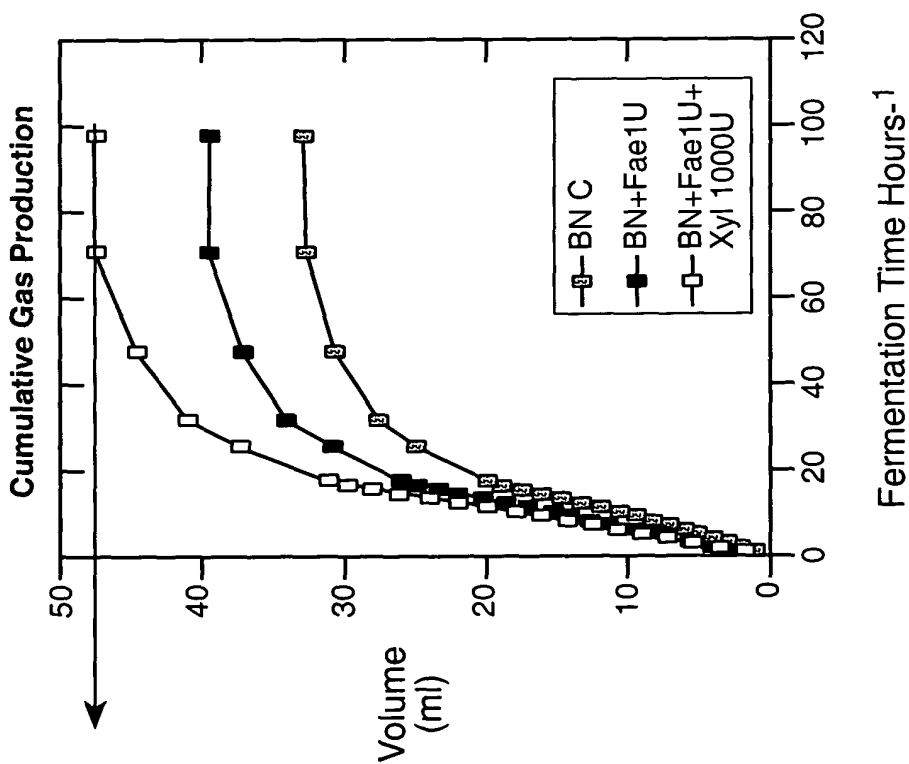


FIG..25A

29 / 154

**In-vitro Fermentation of *Festuca arundinacea* C II Walls
 From C II Cultures Expressing Recombinant FAE1**

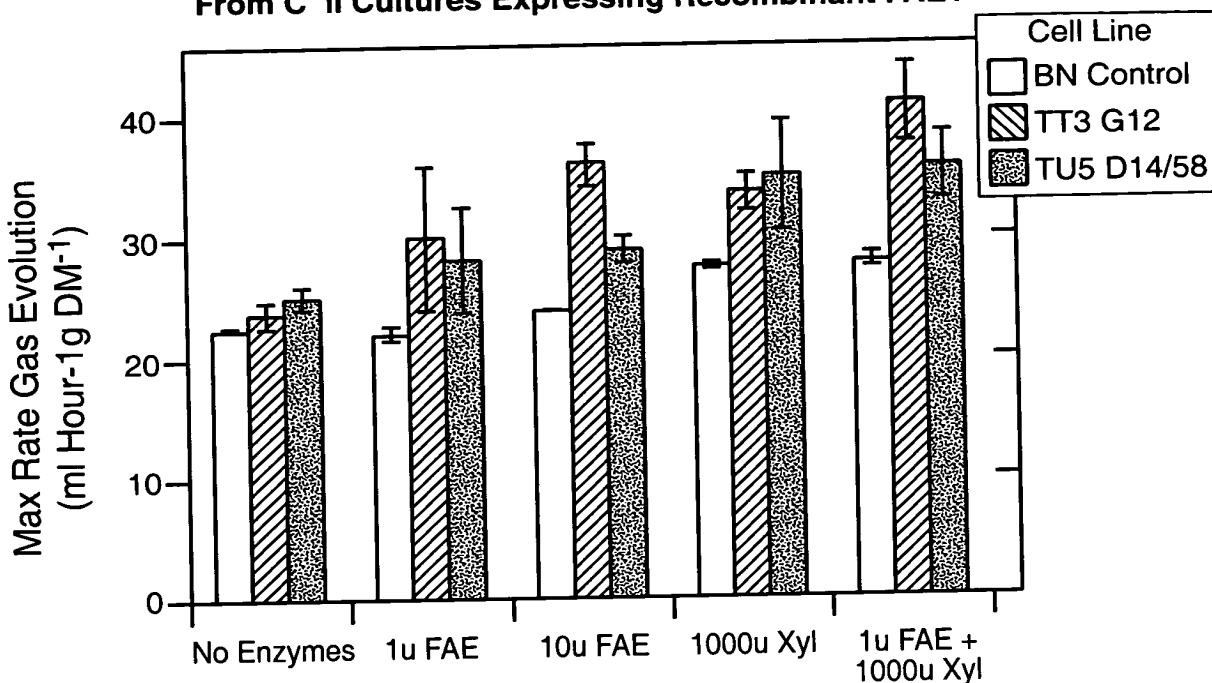


FIG._26A Maximum Rate of Digestion

**In-vitro Fermentation of *Festuca arundinacea* Cell Walls
 From Cell Cultures Expressing Recombinant FAE1**

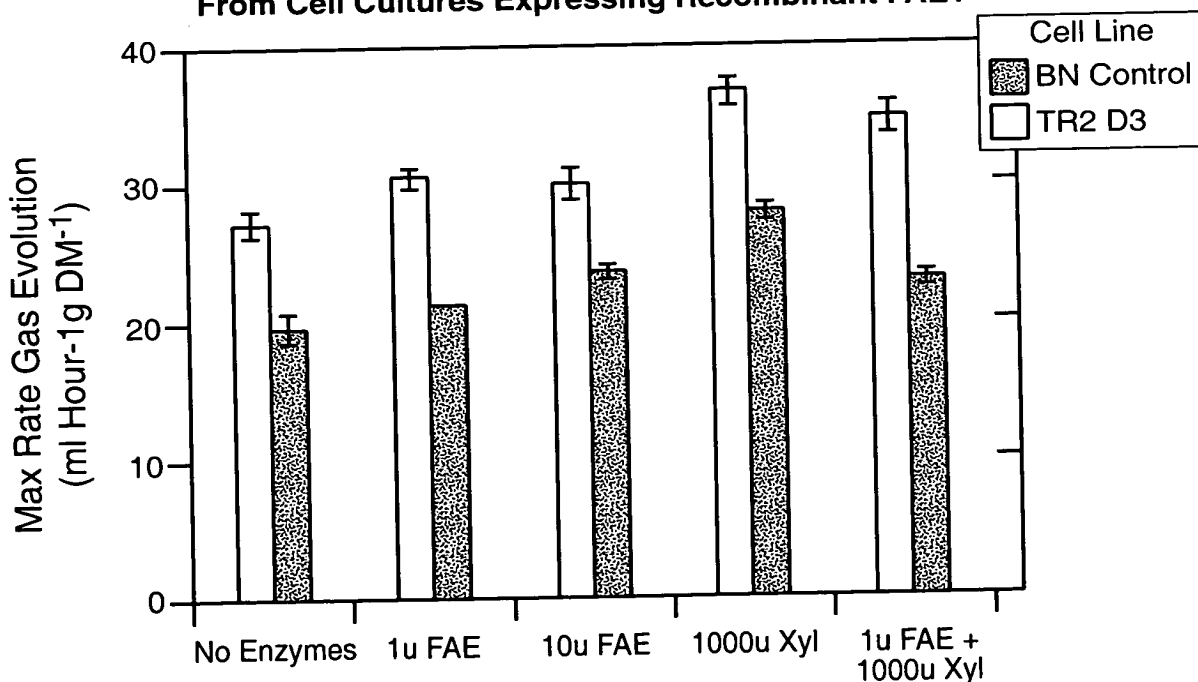


FIG._26B Maximum Rate of Digestion

30 / 154

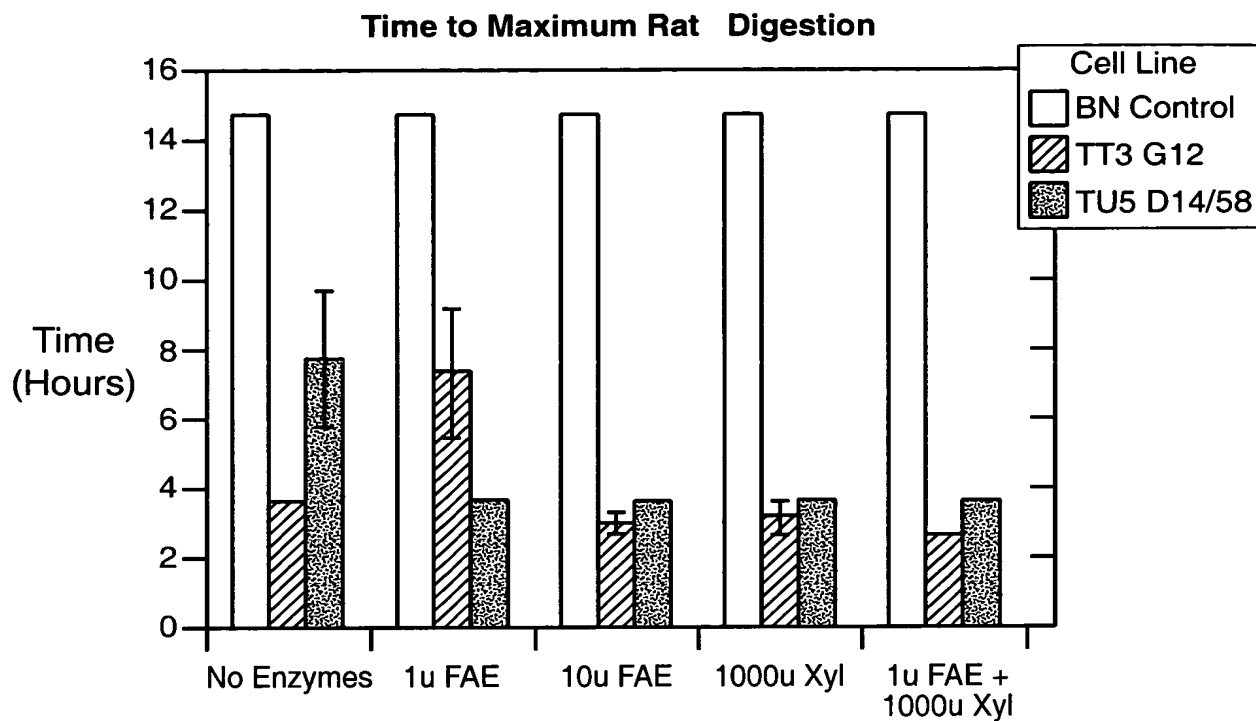


FIG._27A

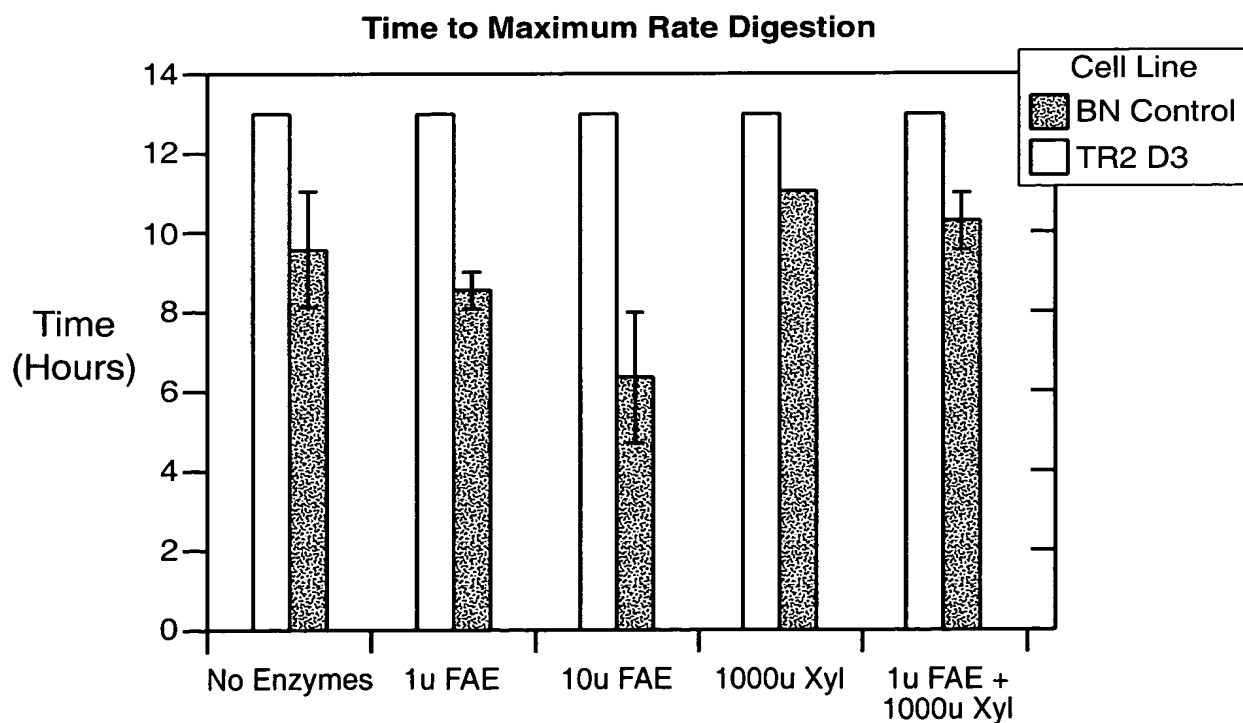


FIG._27B

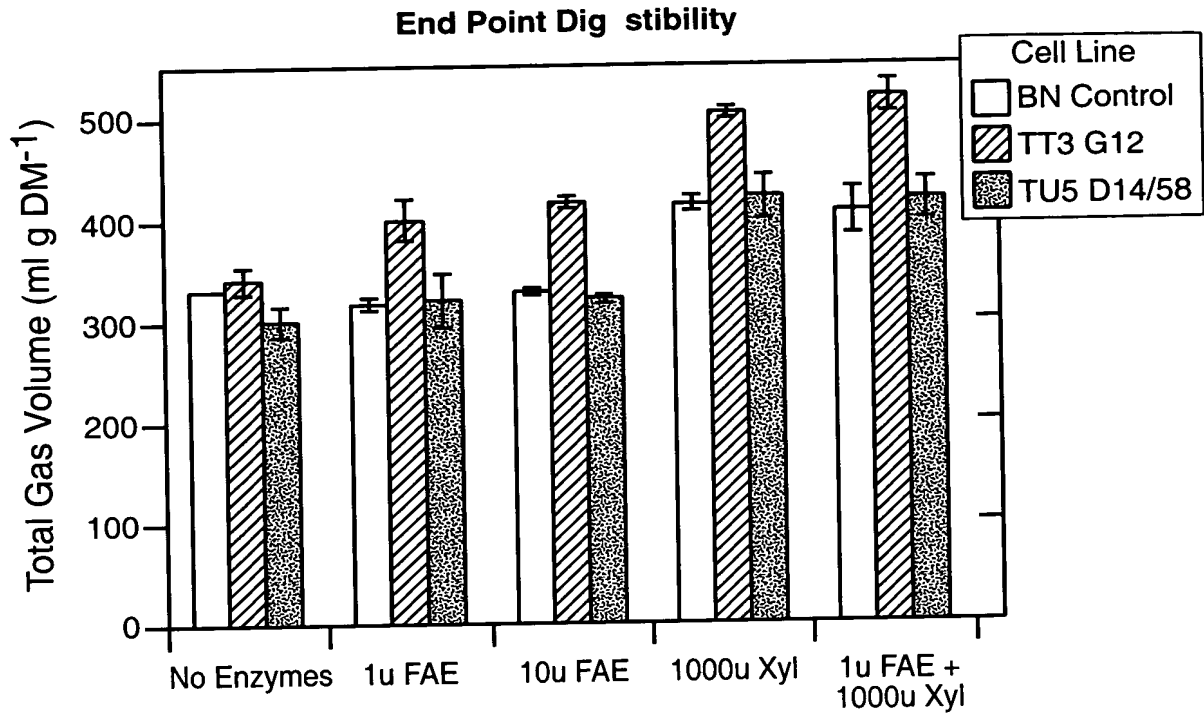


FIG._28A

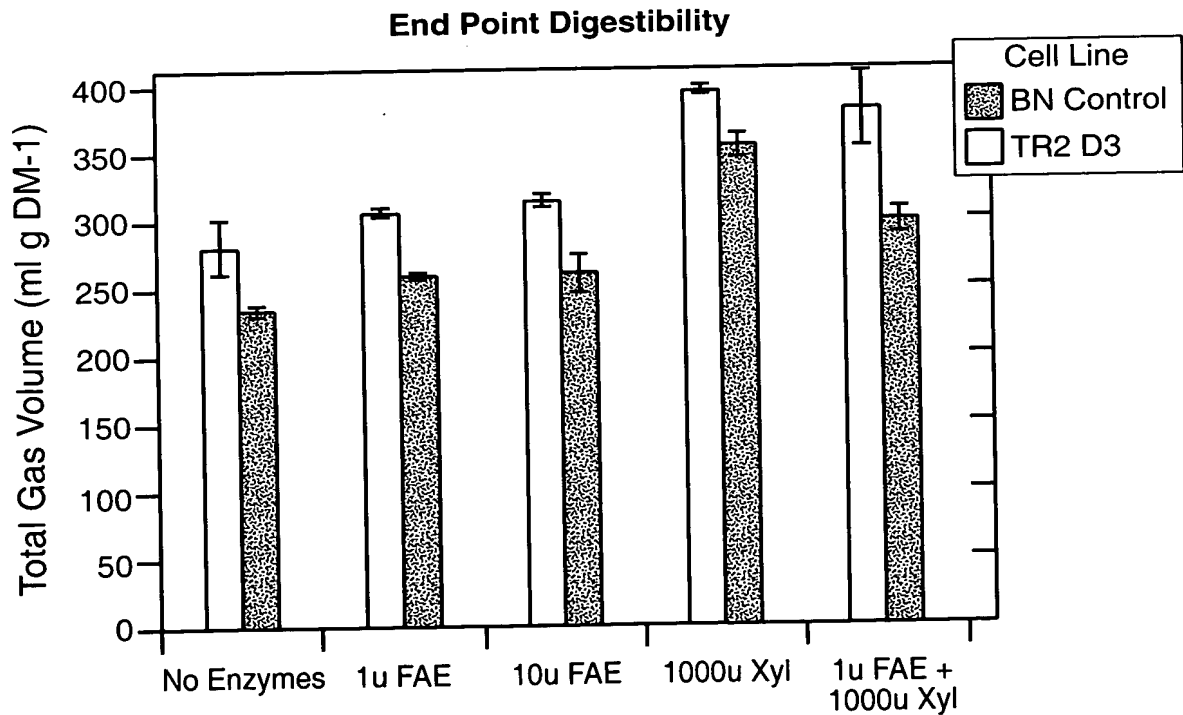


FIG._28B

32 / 154

**Kinetics of FAE Activity by Ferulic Acid Release
 from Cell Wall under Self Digestion in *Festuca arundinacea*
 and Stimulation by Xylanase**

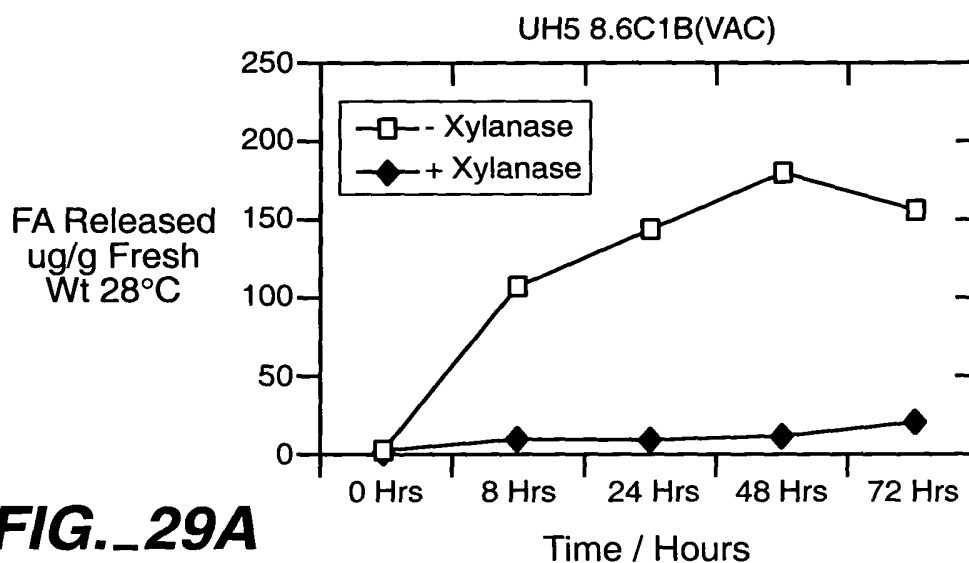


FIG._29A

**Kinetics of FAE Activity by Ferulic Acid Release
 from Cell Wall under Self Digestion in *Festuca arundinacea*
 and Stimulation by Xylanase.**

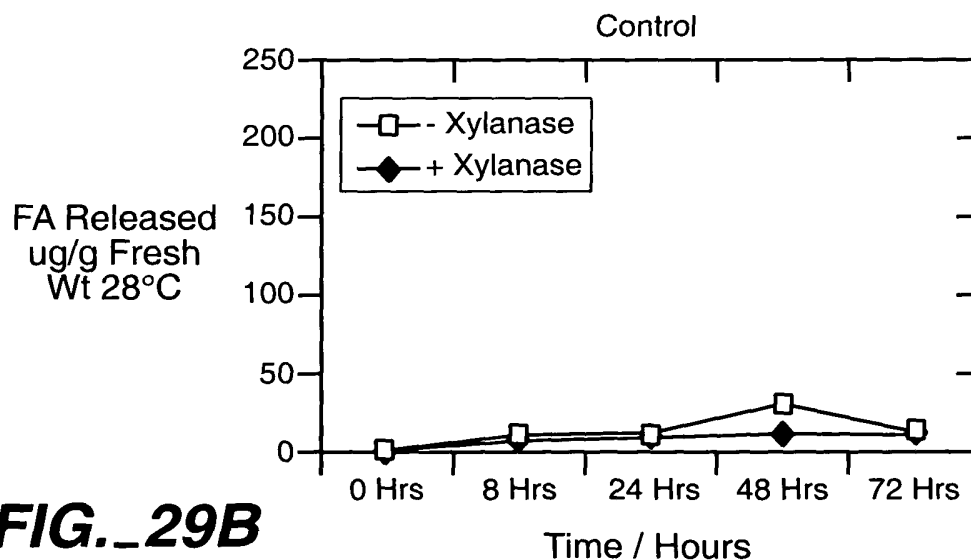


FIG._29B

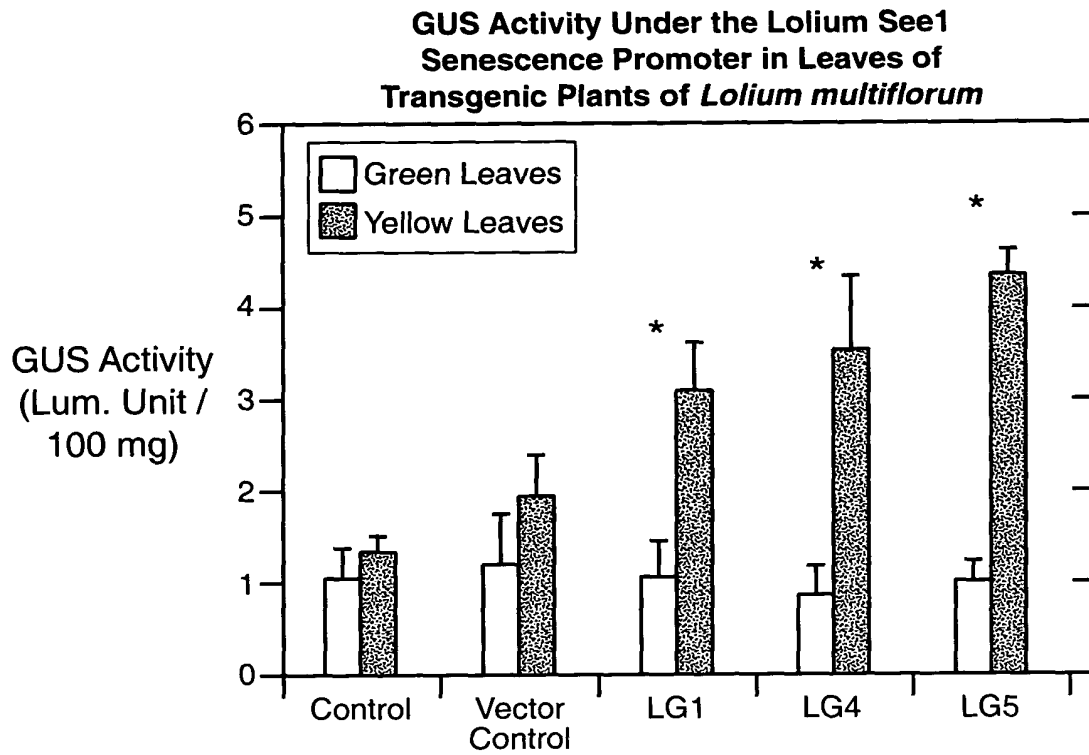


FIG._30

34 / 154

**Release of Monomeric and Dimeric HCAs
on Self Digestion of Leaves of Vacuolar
Targeted FAE Expressing Plants**

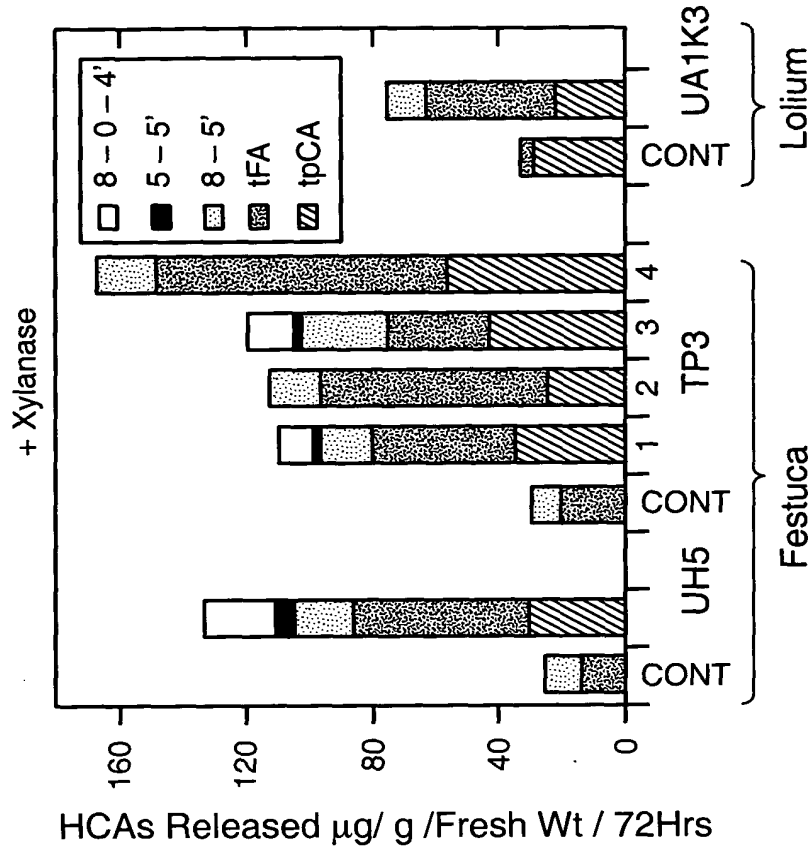


FIG._31B

**Release of Monomeric and Dimeric HCAs
on Self Digestion of Leaves of Vacuolar
Targeted FAE Expressing Plants**

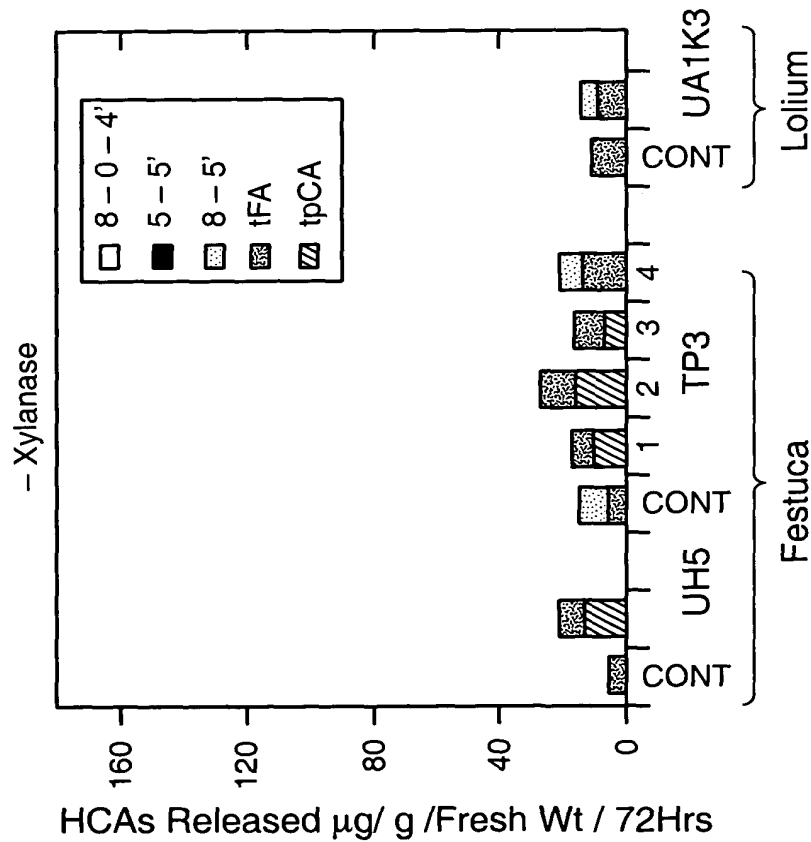


FIG._31A

35 / 154

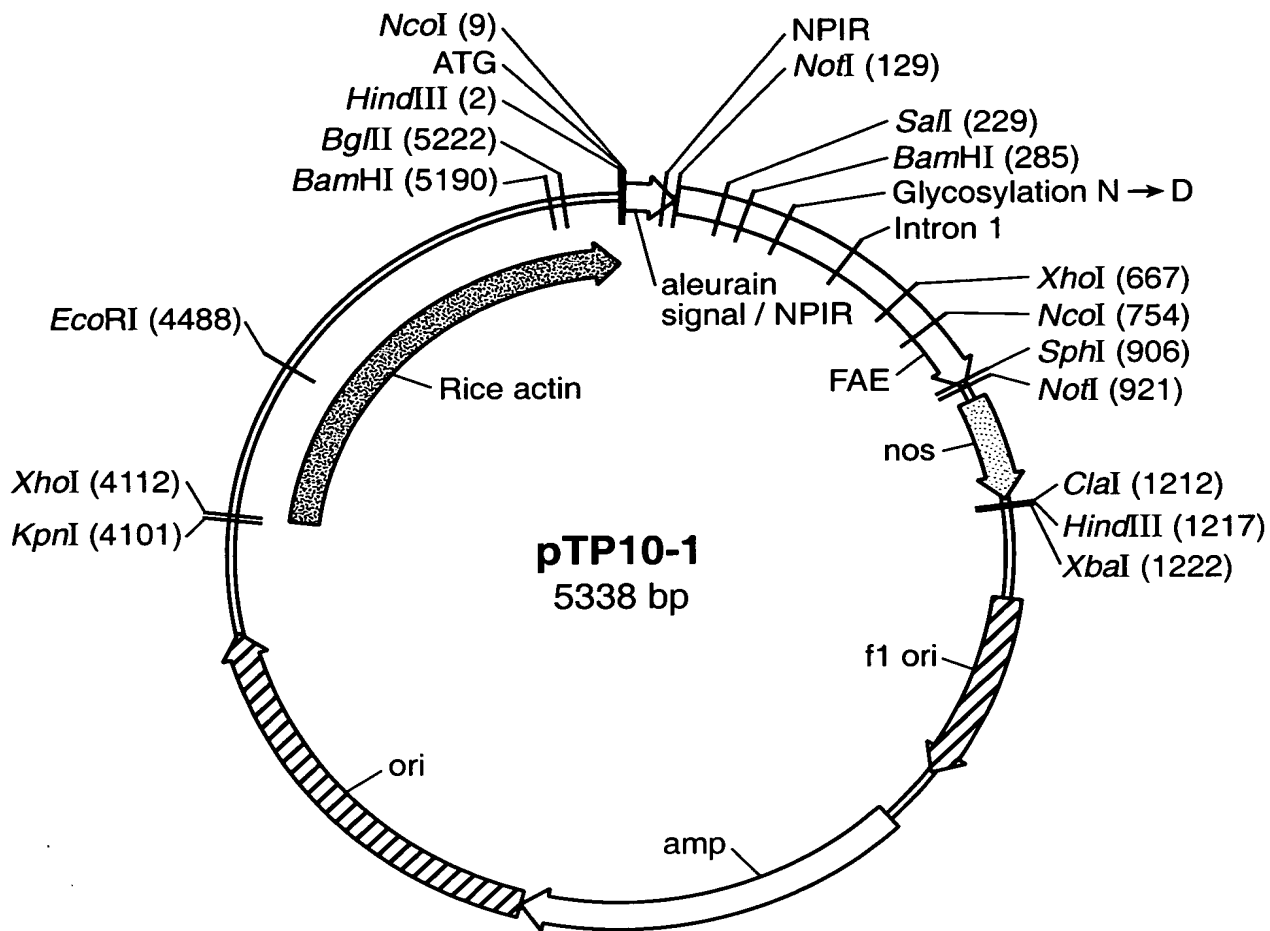


FIG._32A

36 / 154

```

                                NcoI
                                ~~~~~
HindIII
~~~~~
      M   A   H   A   R   V   L   L   L   A   L   A   V   L   A   T   A   A   V   A   V
1  AAGCTTACCA TGGCCACAGC CCGCGTCCTC CTCCTGGCGC TCGCCGTGCT GCCACAGGCC GCCGTGCGCG

                                NPIR
                                ~~~~~
                                NotI
                                ~~~~~
      .   A   S   S   S   F   A   D   S   N   P   I   R   P   V   T   D   R   A   A   A   S   T   .
71  TCGCCTCCTC CTCCTCCTTC GCCGACTCCA ACCCGATCCG GCCCGTCACC GACCGCGCGG CCGCCTCCAC
      .   Q   G   I   S   E   D   L   Y   S   R   L   V   E   M   A   T   I   S   Q   A   A   Y   A
141 GCAGGGCATC TCCGAAGACC TCTACAGCCG TTAGTCGAA ATGGCCACTA TCTCCCAAGC TGCCTACGCC

                                Sali
                                ~~~~~
      D   L   C   N   I   P   S   T   I   I   K   G   E   K   I   Y   N   S   Q   T   D   I   N   G
211 GACCTGTGCA ACATTCCGTC GACTATTATC AAGGGAGAGA AAATTACAA TTCTCAAAC T GACATTAACG

                                BamHI
                                ~~~~~
      .   W   I   L   R   D   D   S   S   K   E   I   I   T   V   F   R   G   T   G   S   D   T   N   .
281 GATGGATCCT CCGCGACGAC AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCAC TGGTA GTGATACGAA

                                Glycosylation
                                ~~~~~
      .   L   Q   L   D   T   D   Y   T   L   T   P   F   D   T   L   P   Q   C   N   G   C   E   V
351 TCTACAAC TC GATACTGACT ACACCCCTCAC GCCTTTCGAC ACCCTACCAC AATGCAACGG TTGTGAAGTA
      H   G   G   Y   Y   I   G   W   V   S   V   Q   D   Q   V   E   S   L   V   K   Q   Q   V   S
421 CACGGTGGAT ATTATATTGG ATGGGTCTCC GTCCAGGACC AAGTCGAGTC GCTTGTCAAA CAGCAGGTTA
      .   Q   Y   P   D   Y   A   L   T   V   T   G   H   X   L   G   A   S   L   A   L   T   A   .
491 GCCAGTATCC GGACTACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG CACTCACTGC
      .   A   Q   L   S   A   T   Y   D   N   I   R   L   Y   T   F   G   E   P   R   S   G   N   Q
561 CGCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC ACCTTCGGCG AACCGCGCAG CGGCAATCAG
```

FIG._32B

37 / 154

```

XhoI
~~~~~
631  GCCTTCGCGT CGTACATGAA CGATGCCTTC CAAGCCTCGA GCCCAGATAC GACGCAGTAT TTCCGGGGTCA

      H A N D G I P N L P P V E Q G Y A H G G V E Y .
701  CTCATGCCAA CGACGGCATC CCAAACCTGC CCCCGGTGGA GCAGGGGTAC GCCCATGGCG GTGTAGAGTA
      W S V D P Y S A Q N T F V C T G D E V Q C C E
771  CTGGAGCGTT GATCCTTACA GCGCCAGAA CACATTTGTC TGCACTGGGG ATGAAAGTGA GTGCTGTGAG

      NcoI
      ~~~~~

      SphI
      ~~~~~

841  GCCCAGGGCG GACAGGGTGT GAATAATGCG CACACGACTT ATTTGGGAT GACGAGCGGC GCATGCACCT

      NotI
      ~~~~~
      KDEL
      ~~~~~
      *
911  GGCCGGTCGC GGCCGCGGAA ACCACTGAAG GATGAGCTGT AAAGAAGCAG ATCGTTCAAA CATTTGGCAA
981  TAAAGTTTCT TAAGATTGAA TCCTGTTGCC GGTCTTGCGA TGATATATCAT ATAAATTTCTG TTGAATTACG
1051 TTAAGCATGT AATAATTAC ATGTAATGCA TGACGTTATT TATGAGATGG GTTTTATGA TTAGAGTCCC
1121 GCAATTATAC ATTTAATACG CGATAGAAA CAAAATATAG CGCGCAAACT AGGATAAATT ATCGCGCGCG

      HindIII
      ~~~~~
      ClaI
      ~~~~~
      XbaI
      ~~~~~
1191 GTGTCATCTA TGTACTAGA TCGATAAGCT TCTAGAGCGG CCGGTGGAGC TCCAATTCCG CCTATAGTGA
1261 GTCGTATTAC GCGCGCTCAC TGGCCGTCGT TTTACAACGT CGTGACTGGG AAAACCCCTGG CGTTACCCAA
1331 CTTAATCGCC TTGCAGCACA TCCCCCTTC GCCAGCTGGC GTAATAGCGA AGAGGCCCGC ACCGATCGCC
1401 CTTCCCAACA GTTGCGCAGC CTGAATGGCG AATGGGACGC GCCCTGTAGC GGCGCATTA GCGCGGCGGG
```

FIG._32C

1471 TGTGGTGGTT ACGCGCAGCG TGACCGCTAC ACTTGCCAGC GCCCTAGCGC CCGCTCCCTT CGCTTCTCTC
1541 CCTTCCTTTC TCGCCACGTT CGCCGGCTTT CCCCCTCAAG CTCTAAATCG GGGCTCCCTT TAGGGTTCC
1611 GATTAGTGC TTTACGGCAC CTCGACCCCA AAAAATTGA TTAGGGTGAT GTTTCACGTA GTGGGCCATC
1681 GCCCTGATAG ACGGTTTTC TATCTCGGTC TATCTTTTG AATTATAAGG GATTTGCGG ATTTCCGCTT
1751 ACTGGAACAA CACTCAACCC AATTAAACAA AATTAAACGC GAATTTTAA TTTCTATAA CATTACAAT
1821 AATTGGTTAA AATGAGCTG CTTTTCGGG AATGTCGCG AAATGCCCTA TTTGTTTAT TTTCTATAA CATTCAAATA
1891 TTAGGTGGCA CATGAGACAA CATGTCGCTT TAATTCCTT TTTTGGCGCA TTTTGCCTTC AAAAGGAAGA GTATGAGTAT
1961 TGATCCGCT GGTGTCGCCC TATATCCCTT TGCTGAAGT CCGTGAACG TTTTCCCAATG ATGAGCACTT TTACATCGAA TCACCCAGAA
2031 TCAACATTTT CCGGTATAT GCTGTAAGT CCCGTATGA ACAGAAAGC ACAGTAACG GAGCACTCG TGGCATGACA GTAGAGAAAT
2101 ACGCTGGTGA AAGTAAAGA GATCCTTGAG AGTTTTCGCC CCGAGAACG TTTTCCCAATG ATGAGCACTT TTACATCGAA GTAGAGAAAT
2171 ACAGCGGTAA GATCCTTGAG GCGGTATAT TGGTGAAGT CCGTGAACG TTTTCCCAATG ATGAGCACTT TTACATCGAA GTAGAGAAAT
2241 GCTATGTGGC CAGATGACT TGGTGAAGT TGGTGAAGT CCGTGAACG TTTTCCCAATG ATGAGCACTT TTACATCGAA GTAGAGAAAT
2311 CAGATGACT TGGTGAAGT TGGTGAAGT CCGTGAACG TTTTCCCAATG ATGAGCACTT TTACATCGAA GTAGAGAAAT
2381 TATGCAGTGC TGCCATAACC ACCGCTTTT TGCACAGCT ACAGTAACG ACAGTAACG GAGCACTCG TGGCATGACA GTAGAGAAAT
2451 GAAGGAGCTA ACCGCTTTT TGCACAGCT ACAGTAACG ACAGTAACG GAGCACTCG TGGCATGACA GTAGAGAAAT
2521 CTGAATGAAG CCATACCAAA TGGCGAATA CTTACTCTAG CGACGAGCT GACACCAACG TTTTCCCAATG ATGAGCACTT TTACATCGAA GTAGAGAAAT
2591 AACTATTAA CTTCTCTGC CTTCTCTGC CTTCTCTGC CTTCTCTGC CTTCTCTGC CTTCTCTGC CTTCTCTGC CTTCTCTGC CTTCTCTGC
2661 AGTGTGAGGA CTTCTCTGC CTTCTCTGC CTTCTCTGC CTTCTCTGC CTTCTCTGC CTTCTCTGC CTTCTCTGC CTTCTCTGC CTTCTCTGC
2731 GAGCGTGGGT CTCGCGGTAT CATGTCAGCA ACTATGGATG AACGAAATAG ACAGTAACG GAGCACTCG TGGCATGACA GTAGAGAAAT
2801 ACACGACGGG GAGTCAGGCA TAACGTCTCA TCTAGGTGAA TCTAGGTGAA TCTAGGTGAA TCTAGGTGAA TCTAGGTGAA TCTAGGTGAA TCTAGGTGAA
2871 TAAGCATTTG TCTAGGTGAA TCTAGGTGAA TCTAGGTGAA TCTAGGTGAA TCTAGGTGAA TCTAGGTGAA TCTAGGTGAA TCTAGGTGAA TCTAGGTGAA
2941 TTTAAAGGA TCTAGGTGAA TCTAGGTGAA TCTAGGTGAA TCTAGGTGAA TCTAGGTGAA TCTAGGTGAA TCTAGGTGAA TCTAGGTGAA TCTAGGTGAA
3011 TCCACTGAGC GTCAGACCCC CAAACAAAAG AGGTAACTGG CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC
3081 CTGCTGCTTG TAGGCCACCA CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC
3151 CTTTTCCTGA AGGTAACTGG CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC
3221 TAGGCCACCA CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC CTTCAAGAAC
3291 TGCTGCGGCT GAACGGGGG TGAGCTATGA GAAAGCGCCA CAGCTTCCCGA CAGCTTCCCGA CAGCTTCCCGA CAGCTTCCCGA CAGCTTCCCGA CAGCTTCCCGA
3361 CGGTCGGGCT GAACGGGGG TGAGCTATGA GAAAGCGCCA CAGCTTCCCGA CAGCTTCCCGA CAGCTTCCCGA CAGCTTCCCGA CAGCTTCCCGA CAGCTTCCCGA
3431 ACCTACAGCG GAAACAGGAG AGCGACGCGT CTTTTCACGGT TCGTGGGCTT TCGTGGGCTT TCGTGGGCTT TCGTGGGCTT TCGTGGGCTT TCGTGGGCTT
3501 CGGCAGGGTC GCCACCTCTG ACTTGAGCGT TTTTTCACGGT TCGTGGGCTT TCGTGGGCTT TCGTGGGCTT TCGTGGGCTT TCGTGGGCTT TCGTGGGCTT
3571 GTCGGGTTTC CCAACGCGGC CCGATTTCTG TCGTGGGCTT TCGTGGGCTT TCGTGGGCTT TCGTGGGCTT TCGTGGGCTT TCGTGGGCTT TCGTGGGCTT
3641 AAAACGCGCG CCGATTTCTG TGGATAACCG TATTAACCGC TTTGAGTGAG CTGATACCGC TTTGAGTGAG CTGATACCGC TTTGAGTGAG CTGATACCGC
3711 TCGGTTATCC CCGATTTCTG TGGATAACCG TATTAACCGC TTTGAGTGAG CTGATACCGC TTTGAGTGAG CTGATACCGC TTTGAGTGAG CTGATACCGC
3781 CGAACGACCG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG AAGAGCGCCC AATACGCAAA AATACGCAAA AATACGCAAA AATACGCAAA
3851 CCGCGCGTTG GCCGATTCTA TAATGCAGCT GGCACGACAG GTTTCCTCCGAC TGGAAAGCGG GCAGTGAGCG

FIG._32D

39 / 154

```
3921 CAACGCAATT AATGTGAGTT AGCTCACTCA TTAGGCACCC CAGGCTTTAC ACTTTATGCT TCCGGCTCGT
3991 ATGTTGTGTG GAATTGTGAG CGGATAACAA TTTCACACAG GAAACAGCTA TGACCATGAT TACGCCAAGC
      ~~~~~
      KpnI
4061 GCGCAATTAA CCTCACTAA AGGGAACAAA AGCTGGGTAC CGGGCCCCC CTCGAGGTCA TTCATATGCT
4131 TGAGAAGAGA GTCGGGATAG TCCAAAATAA AACAAAGGTA AGATTACCTG GTCAAAAGTG AAAACATCAG
4201 TTAAAAGGTG GTATAAGTAA AATATCGGTA ATAAAAGGTG GCCCAAAGTG AAATTACTC TTTTCTACTA
4271 TTATAAAAAA TGAGGATGTT TTGTCGGTAC TTTGATACGT CATTTTGTG TGAATTGGTT TTTAAGTTTA
4341 TTCGCGAATT GGAATGCAAT ATCTGTATTT GAGTCGGTTT TTAAGTTCTG TGCTTTTGTG AATACAGAGG
4411 GATTGTGATA AGAATATCTT TTAATAAACCAT CATATGCTAA TTTGACATAA TTTTGTGAGAA AAATATATAT
      ~~~~~
      EcorI
4481 TCAGGCGGAAT TCCACAATGA ACAATAATAA GATTAAAATA GCTTGCCCCC GTTGCAGCGA TGGGTATTTT
4551 TTCTAGTAAA ATAAAAGATA AACTTAGACT CAAAACATTT ACAAAAACAA CCCCATAAGT CCTAAAGCCC
4621 AAAGTGCTAT GCACGATCCA TAGCAAGCCC AGCCCAACCC AACCCAAACC AACCCACCCC AGTGCAGCCA
4691 ACTGGCAAAAT AGTCTCCACC CCGGCACTA TCACCGTGAG TTGTCCGCAC CACCGCACGT CTCGCAGCCA
4761 AAAAAAATAA AAGAAAGAAA AAAAAGAAA AGAAAACAG CAGGTGGGTC CGGGTCGTGG GGGCCGGAAA
4831 AGCGAGGAGG ATCGCGAGCA GCGACGAGGC CCGGCCCTCC CTCCGCTTCC AAAGAAACGC CCCCCATCGC
4901 CACTATATAC ATACCCCCC CTCTCTCTCC ATCCCCCCTAA CCTACCCACC ACCACCAACA CCACCTCCTC
4971 CCCCCTCGCT GCCGGACGAC GAGCTCCTCC CCCCCTCCCC TCCGCCGCCG CCGGTAACCA CCCCCCCCCC
5041 CTCCTCTTTC TTTCTCCGTT TTTTCTTTCG TCTCGGTCCT GATCTTTGGC CTGGGTAGTT TGGGTGGGCG
5111 AGAGCGGCTT CGTCGCCCCAG ATCGGTGCGC GGGAGGGGCG GGATCTCGCG GCTGGCGTCT CCGGGCGTGA
      ~~~~~
      BamHI
5181 GTCGGCCCCG ATCCTCGCGG GGAATGGGGC TCCTCGGATGT AGATCTTCTT TCCTTCTTCT TTTTGTGGTA
5251 GAATTGGAAT CCTCAGCAT TGTTCATCGG TAGTTTCTCT TTTTCATGAT TTTCATGAAAT GCAGCCTCGT
5321 GCGGAGCTTT TTTGTAGC
```

FIG._32E

40 / 154

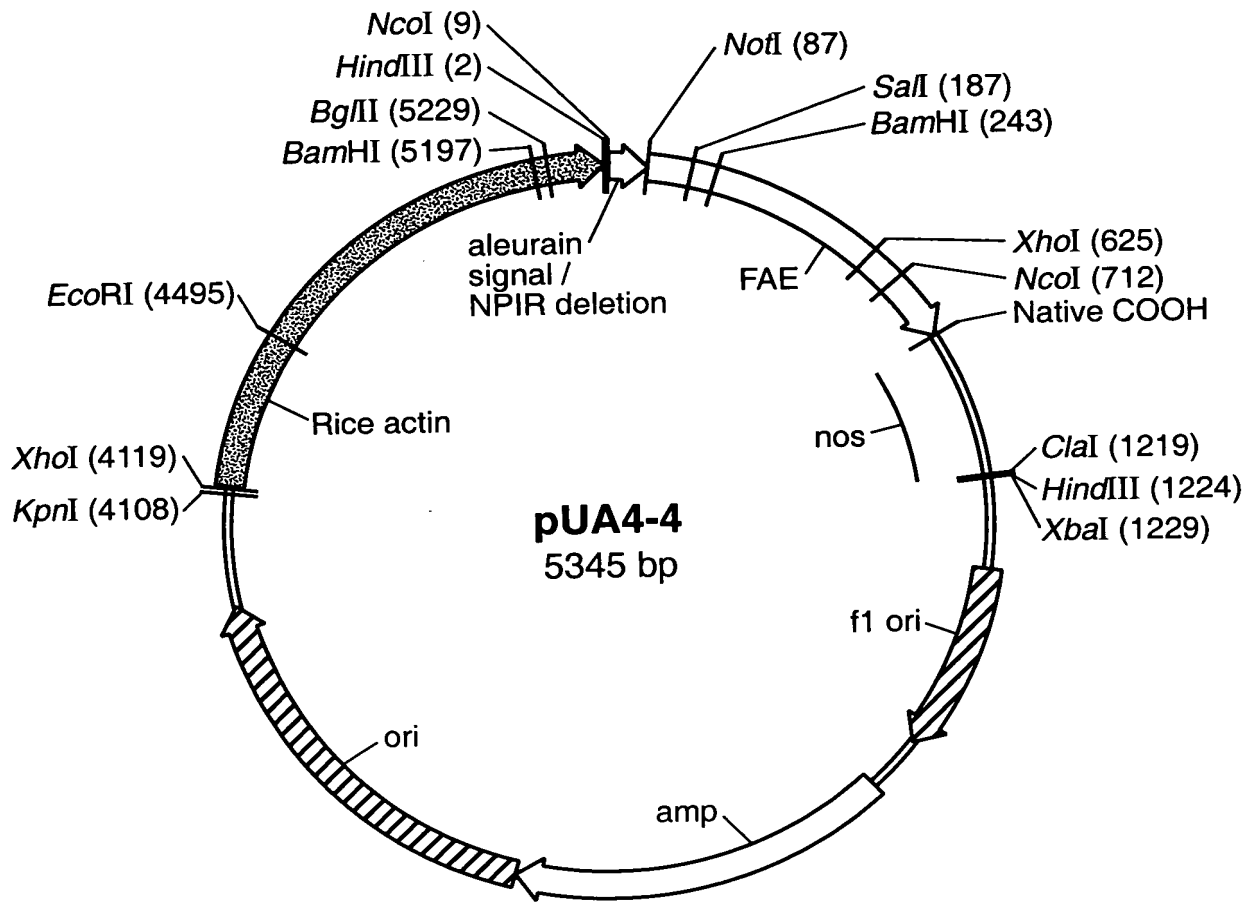


FIG._33A

41 / 154

```

                                NcoI
                                ~~~~~
HindIII
~~~~~
      M   A   H   A   R   V   L   L   L   A   L   A   V   L   A   T   A   A   V   A   V
1  AAGCTTACCA TGGCCACACGC CCGCGTCTCTC CTCTCTGGCGC TCGCCGTGCT GGCACAGGCC GCCGTGCGCG
                                NotI
                                ~~~~~
      .   A   S   S   R   A   A   A   S   T   Q   G   I   S   E   D   L   Y   S   R   L   V   E   M   .
71 TCGCCTCCTC CCGCGCGGCC GCCTCCACGC AGGCATCTC CGAAGACCTC TACAGCCGTT TAGTCGAAAT
                                Sali
                                ~~~~~
      .   A   T   I   S   Q   A   A   Y   A   D   L   C   N   I   P   S   T   I   I   K   G   E   K
141 GGCCACTATC TCCCAAGCTG CCTACGCCGA CCTGTGCAAC ATTCCGTGCA CTATTATCAA GGGAGAGAAA
                                BamHI
                                ~~~~~
      I   Y   N   S   Q   T   D   I   N   G   W   I   L   R   D   D   S   S   K   E   I   I   T   V
211 ATTACAATT CTCAAACTGA CATTAACGGA TGGATCCTCC GCGACGACAG CAGCAAAGAA ATAATCACCG
      .   F   R   G   T   G   S   D   T   N   L   Q   L   D   T   N   Y   T   L   T   P   F   D   T   .
281 TCTTCCGTGG CACTGGTAGT GATACGAATC TACAACCTGA TACTAACTAC ACCCTCACGC CTTTCGACAC
      .   L   P   Q   C   N   G   C   E   V   H   G   G   Y   Y   I   G   W   V   S   V   Q   D   Q
351 CCTACCACAA TGCAACGGTT GTGAAGTACA CGGTGGATAT TATATTGGAT GGGTCTCCGT CCAGGACCAA
      V   E   S   L   V   K   Q   Q   V   S   Q   Y   P   D   Y   A   L   T   V   T   G   H   X   L
421 GTCGAGTCGC TTGTCAAACA GCAGGTTAGC CAGTATCCGG ACTACGCGCT GACCGTGACC GGCCACKCCC
      .   G   A   S   L   A   A   L   T   A   A   Q   L   S   A   T   Y   D   N   I   R   L   Y   T   .
491 TCGGCGCCTC CCTGGCGGCA CTCACTGCGC CCCAGCTGTC TCGGACATAC GACAACATCC GCCTGTACAC
                                XhoI
                                ~~~~~
      .   F   G   E   P   R   S   G   N   Q   A   F   A   S   Y   M   N   D   A   F   Q   A   S   S
561 CTTGGCGGAA CCGCGCAGCG GCAATCAGGC CTTCGCGTCG TACATGAACG ATGCCCTTCCA AGCCTCGAGC
      P   D   T   T   Q   Y   F   R   V   T   H   A   N   D   G   I   P   N   L   P   P   V   E   Q
631 CCAGATACGA CGCAGTATTT CCGGGTCACT CATGCCAACG ACGGCATCCC AAACCTGCC CCGGTGGAGC
                                NcoI
                                ~~~~~
      .   G   Y   A   H   G   G   V   E   Y   W   S   V   D   P   Y   S   A   Q   N   T   F   V   C   .

```

FIG._33B

```

701 AGGGGTACGC CCATGGCGGT GTAGAGTACT GGAGCGTTGA TCCTTACAGC GCCCAGAACA CATTGTCTCG
    . T G D E V Q C C E A Q G G Q G V N A H T T Y
771 CACTGGGGAT GAAGTGCAGT GCTGTGAGGC CCAGGGCGGA CAGGTGTGA ATAATGCGCA CACGACTTAT
    F G M T S G A C T W *
841 TTTGGGATGA CGAGCGGAGC CTGTACATGG TGATCAGTCA TTTCAGCCTC CCCGAGTGTA CCAGGAAAGA
911 TGGATGTCCT GGAGAGGGGG CCGCGTAACC ACTGAAGGAT GAGCTGTAAA GAAAGCAGATC GTTCAAAACAT
981 TTGGCAATAA AGTTTCTTAA GATTGAATCC TGTTCGCGGT CTTGCGATGA TTATCATATA ATTTCTGTTG
1051 AATTACGTTA AGCATGTAAT AATTAAACATG TAATGCATGA CGTTATTTAT GAGATGGGTT TTTATGATTA
1121 GAGTCCCGCA ATTATACATT TAATACGCGA TAGAAAACAA AATATAGCGC GCAAACTAGG ATAAATTATC

                                HindIII
                                ~~~~~
                                ClaI      XbaI
                                ~~~~~
1191 GCGCGCGGTG TCATCTATGT TACTAGATCG ATAAGCTTCT AGAGCGGCGG GTGGAGCTCC AATTCGCCCT
1261 ATAGTGAGTC GTATTACGCG CGTCACTGG CCGTCGTTT ACAACGTCGT GACTGGGAAA ACCCTGGCGT
1331 TACCCAACTT AATCGCCTTG CAGCACATCC CCTTTTCGCC AGCTGGCGTA ATAGCGAAGA GGCCCGCACC
1401 GATCGCCCTT CCCAACAGTT GCGACGCTG AATGGCGAAT GGGACGCGCC CTGTAGCGGC GCATTAAGCG
1471 CGGCGGGTGT GGTGGTTACG CGCAGCGTGA CCGTACACT TGCCAGCGCC CTAGCGCCCG CTCCTTTTCG
1541 TTTCTTCCCT TCCTTTCTCG CCACGTTCCG CGGCTTTCCC CGTCAAGCTC TAAATCGGGG GCTCCCTTTA
1611 GGGTTCGGAT TTAGTGCTTT ACGGCACCTC GACCCCAAAA AACTTGATTA GGGTGATGGT TCACGTAAGT
1681 GGCCATCGCC CTGATAGACG GTTTTTCGCC CTTTGACGTT GGAGTCCACG TCTTTTAAATA GTGGACTCTT
1751 GTTCCAAACT GGAACAAAC GAAACCTAT CTGCGTCTAT TCTTTTGAAT TATAAGGGAT TTTGCCGATT
1821 TCGGCCTATT GGTAAAAAA TGAGCTGATT TAACAAAAAT TTAACGCGAA TTTTAAACAAA ATATTAACGC
1891 TTACAAATTA GTGGCACTT TTGCGGGAAA TGTCGCGCGA ACCCTATTT GTTTATTTTT CTAAATACAT
1961 TCAATATATG ATCCGCTCAT GAGACAATAA CCTGATAAAA TGCTTCAATA ATATGAAAA AGGAAGAGTA
2031 TGAGTATTCA ACATTTCCTG GTCGCCCTTA TTCCCTTTTT TCGGGCATTT TGCCCTCCCTG TTTTGTCTCA
2101 CCCAGAAACG CTGGTGAAG TAAAGATGC TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAACTG
2171 GATCTCAACA GCGGTAAGAT CCTTGAGAGT TTTGCGCCCG AAGAACGTTT TCCAATGATG AGCACTTTAA
2241 AAGTTCTGCT ATGTGGCGG GTATTATCCC GTATTGACGC CGGGCAAGAG CAACTCGGTC GCCGCATACA
2311 CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTCACA GAAAGCATC CTGCGGCCAA CTTACTTCTG ACAACGATCG
2381 AGAGAATTAT GCAGTGCTGC CATAACCATG AGTGATAACA CTGCGGCCAA GATCATGTA ACTCGCTTG ATCGTTGGGA
2451 GAGGACCGAA GGAGCTAACC GCTTTTTCG ACAACATGGG GGATCATGTA ACCCGCTTG ACTCGTTGGG
2521 ACCGGAGCTG AATGAAGCCA TACCAAACGA CGAGCGTGAC ACCCGATGC CTGTAGCAAT GGCAACAACG
2591 TTGCGCAAAC TATTAACTGG CGAACTACTT ACTCTAGCTT CCCGGCAACA ATTAATAGAC TGGATGGAGG
2661 CGGATAAAGT TGCAGGACCA CTTCCTGCGCT CGGCCCTTCC GGCTGGCTGG TTTTATTGCTG ATAAATCTTG

```

FIG._33C

```

2731 AGCCGGTGAG CGTGGGTCTC GCGGTATCAT TGCAGCACTG GGGCCAGATG GTAAAGCCCTC CCGTATCGTA
2801 GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC GAAATAGACA GATCGCTGAG ATAGGTGCCT
2871 CACTGATTAA GCATTGGTAA CTGTCAGACC AAGTTTACTC ATATATACTT TAGATTGATT TAAAACTTCA
2941 TTTTAAATTT AAAAGGATCT AAGTGAAGAT GAAAGATCA AAGGATCTTC AATCTCATGA CCAAAATCCC TTTTCTCTGC
3011 TTTTCGTTCC ACTGAGCGTC AGACCCCGTA AGACCCGTA CAGCGTACC AGCGTGGTT TGTTCGCCG ATCAAGAGCT
3081 GCSTAATCTG CTGCTTGCAA ACAAAAAC TAACCTGGCTT CAGCAGAGCG CAGATACCAA ATAGTGTCTT TCTAGTGTAG
3151 ACCAACTCTT TTTCCGAAGG GCCACCACTT CAAGAACTCT GATGACCCG GTCTTACCAG GTTGGACTCA AGACGATAGT TACCGGATAA
3221 CCGTAGTTAG GACCACTCTT TGCCAGTGGC GATAAGTCTG GTGACACACG CCCAGCTGG AGCGAACGAG GAGAAAGGCG GACAGGTATC
3291 CAGTGGCTGC TGCCAGTGGC GATAGTCTG GATAGTCTG GTGACACACG CCCAGCTGG AGCGAACGAG GAGAAAGGCG GACAGGTATC
3361 GCGCAGCGG TCGGGCTGAA GCTATGAGAA AGCGCCACAG AGCGCTGG GTTCCGAAG GCTTCCAGG GGTATCTTTA GGTATCTTTA
3431 CTGAGATACC TACAGCGTGA CAGGTCGGA CAGGTCGGA ACCTCTGACT TGACCGTCC TTAACCGTAT TACCGCTTTT CTGCGCTCG
3501 CGGTAAGCGG CAGGTCGGA CAGGTCGGA ACCTCTGACT TGACCGTCC TTAACCGTAT TACCGCTTTT CTGCGCTCG
3571 TAGTCCTGTC GGTTCGCGC ACGCCAGCAA CGCGGCTCTT GATTCGTGG ATACCGCTAT TACCGCTTTT CTGCGCTCG
3641 CTATGGAAA ACGCCAGCAA CGCGGCTCTT GATTCGTGG ATACCGCTAT TACCGCTTTT CTGCGCTCG
3711 TCTTCTCTGC GTTATCCCTT GATTCGTGG ATACCGCTAT TACCGCTTTT CTGCGCTCG
3781 CCGCAGCCGA ACGACCGAGC CAGGTCGGA CAGGTCGGA ACCTCTGACT TGACCGTCC TTAACCGTAT TACCGCTTTT CTGCGCTCG
3851 CCTCTCCCCG CGCGTTGGCC GATTCATTA GATTCATTA TGCAGCTGGC AGCAGCGGAG GAGCGCCAA AGCGCCAA ACGCAAAACCG
3921 GTGAGCGCAA CGCAATTAAT GTGAGTTAG TCACTCATTA GGCACCCAG GCTTTACACT TTTATGCTTCC
3991 GGCTCGTATG TTGTGTGGAA TTGTGAGCGG ATAACAATTT CACACAGGAA ACAGCTATGA CCATGATTAC

4061 GCCAAGCGCG CAATTAAACC TCACATAAAG GAACAAAAGC TGGGTACCGG GCCCCCCCTC GAGGTCAATC
4131 ATATGCTTGA GAAGAGAGTC GGGATAGTCC AAAATAAATC AAAGGTAAGA TTACCTGGTC AAAAGTGAAA
4201 ACATCAGTTA AAAGGTGGTA TAAAGTAAAT ATCGGTAATA AAAGGTGGCC CAAAGTGAAA TTTACTCTTT
4271 TCTACTATTA TAAAAATTGA GATGTTTTG TCGGTACTTT GATACGTCTT TTTTGTATGA ATTGGTTTTT
4341 AAGTTTATTC GCGATTGGA AATGCATATC TGTATTTAG TCGGTTTTTA AGTTCGTGTC TTTTGTAAAT
4411 ACAGAGGGAT TTGTATAAGA AATATCTTTA AAAAACCCAT ATGCTAATTT GACATAAATTT TTGAGAAAAA

4481 TATATATTCA GCGGAATTCC ACAATGAACA ATAATAAGAT TAAAATAGCT TGCCCCCGTT GCAGCGATGG
4551 GTATTTTTC TAGTAAAAA TAAGATAAAC TTAGACTCAA AACATTTACA AAAACAACCC CTAAGTCTCT
4621 AAAGCCCAA GTGCTATGCA CGATCCATAG CAAGCCAGC CCAACCCAA CCAACCCAGT
4691 GCAGCCAACT GGCAAAATAG CTCCACCCCC GGCACATATCA CCGTGAGTTG TCCGCACCCAC CGCACGTCCTC
4761 GCAGCCCAA AAAAAAAAG AAAAAAAA AAGAAAAA ACAGAGCCAG GTGGTCCCG GTCGTGGGGG
4831 CCGGAAAAGC GAGGAGGATC GCGAGCAGCG ACAGGGCCCG GCCCTCCCTC CGCTTCCAAA GAAACGCCCC

```

FIG.-33D

```
4901 CCATCGCCAC TATATACATA CCCCCCCTC TCCTCCCATC CCCCACACC TACCACCACC ACCACCACCA
4971 CCTCCTCCCC CCTCGCTGCC GGACGACGAG CTCTCTCCCC CTCCCCCTCC GCCGCCGCCG GTAACCACCC
5041 CGCCCTCTC CTCTTTCTTT CTCCGTTTTT TTTTTCGTCT CCGTCTCGAT CTTTGGCCCTT GGTAGTTTGG
5111 GTGGGCGAGA GCGGCTTCGT CGCCAGATC GGTGCGCGG GGGGCGGGA TCTCGCGGCT GCGTCTCTCCG
                                     BamHI
                                     ~~~~~
5181 GCGGTGAGTC GGCCCGGATC CTCGCGGGA ATGGGCTCT CGGATGTAGA TCTTCTTTCT TTCTTCTTTT
5251 TGTGGTAGAA TTTGAATCCC TCAGCATTTG TCATCGGTAG TTTTCTTTT CATGATTGT GACAAATGCA
5321 GCGTCGTGCG GAGCTTTT GTAGC
```

FIG._33E

45 / 154

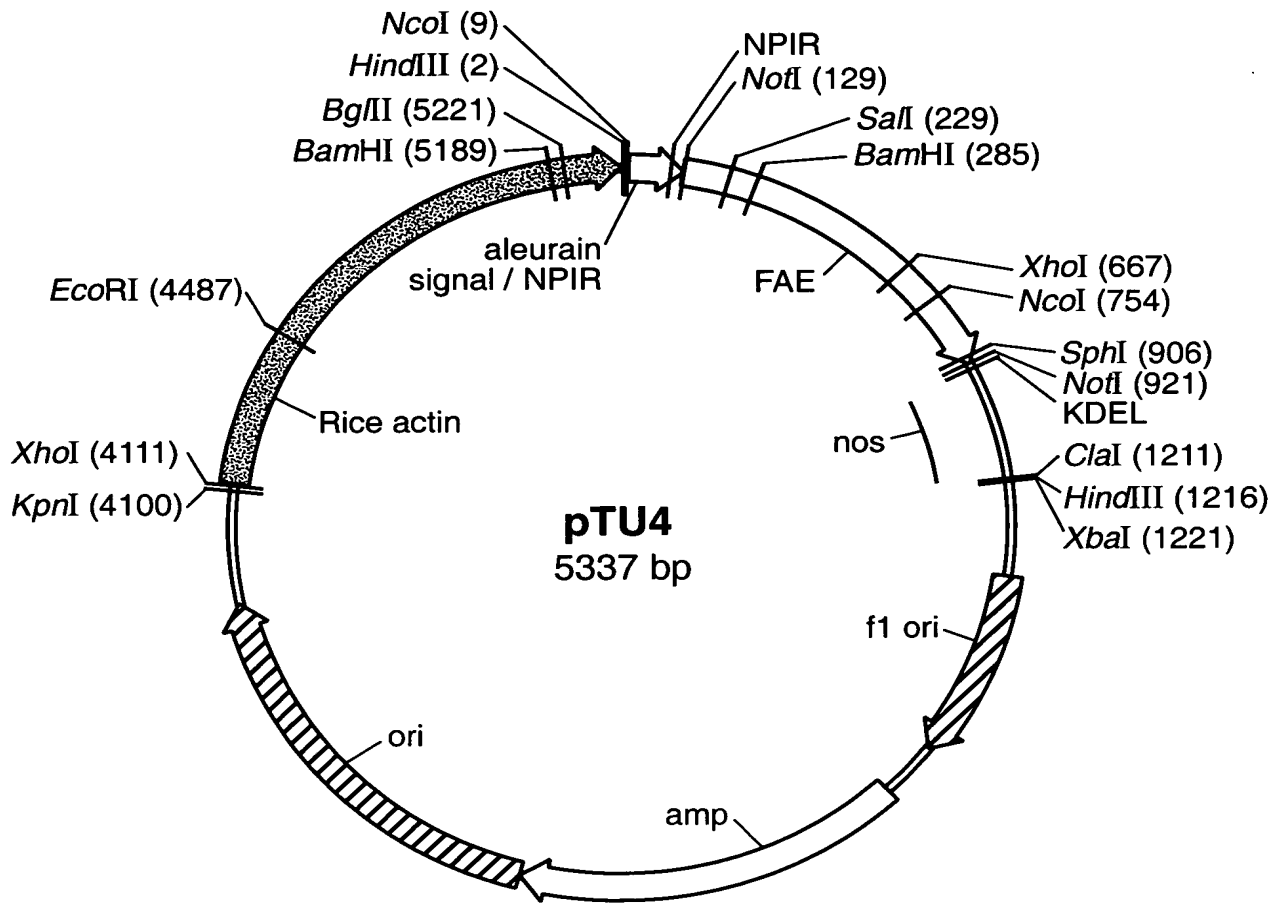


FIG. 34A

46 / 154

```

      NcoI
      ~~~~~
HindIII
~~~~~
1  AAGCTTACCA TGGCCCACGC CCGCGTCTCTC CTCCCTGGCGC TCGCCGTGCT GCCACGGCC GCCGTGCGCC
      NotI
      ~~~~~
      . A S S S S F A D S N P I R P V T D R A A A S T .
71  TCGCCTCCTC CTCCCTCCTTC GCCGACTCCA ACCCGATCCG GCCCGTCACC GACCGCGCGG CCGCCTCCAC
      . Q G I S E D L Y S R L V E M A T I S Q A A Y A
141 GCAGGGCATC TCCGAAGACC TCTACAGCCG TTAGTCGAA ATGGCCACTA TCTCCCAAGC TGCCTACGCC
      Sali
      ~~~~~
      D L C N I P S T I I K G E K I Y N S Q T D I N G
211 GACCTGTGCA ACATTCCGTC GACTATTATC AAGGGAGAGA AAATTACAA TTCTCAAAC T GACATTAAACG
      BamHI
      ~~~~~
      . W I L R D D S S K E I I T V F R G T G S D T N .
281 GATGGATCCT CCGCGACGAC AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCAC TGGTA GTGATACGAA
      . L Q L D T N Y T L T P F D T L P Q C N G C E V
351 TCTACAATC GATACTAAT ACACCTCACC GCCTTTCGAC ACCCTACCAC AATGCAACGG TTGTGAAGTA
      H G G Y Y I G W V S V Q D Q V E S L V K Q Q V S
421 CACGGTGGAT ATTATATTGG ATGGGTCTCC GTCCAGGACC AAGTCGAGTC GCTTGTCAA CAGCAGGTTA
      . Q Y P D Y A L T V T G H X L G A S L A A L T A .
491 GCCAGTATCC GGAATACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG CACTCAGTGC
      . A Q L S A T Y D N I R L Y T F G E P R S G N Q
561 CGCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC ACCTTCGGCG AACCGCGCAG CGGCAATCAG
      XhoI
      ~~~~~
      A F A S Y M N D A F Q A S S P D T T Q Y F R V T
631 GCCTTCGCGT CGTACATGAA CGATGCCTTC CAAGCCTCGA GCCCAGATAC GACGCAGTAT TTCCGGGGTCA
      NcoI
      ~~~~~
      . H A N D G I P N L P P V E Q G Y A H G G V E Y .

```

FIG._34B

47 / 154

```

701 CTCATGCCAA CGACGGCATC CCAAACCTGC CCCCAGGTGA GCAGGGGTAC GCCCATGGCG GTGTAGAGTA
    . W S V D P Y S A Q N T F V C T G D E V Q C C E
771 CTGGAGCGTT GATCCTTACA GCGCCAGAA CACATTGTGTC TGCACCTGGG ATGAAGTGCA GTGCTGTGAG
    SphI
    ~~~~
    A Q G G Q G V N N A H T T Y F G M T S G A C T W
841 GCCCAGGCG GACAGGGTGT GAATAATGCG CACACGACTT ATTTTGGGAT GACGAGCGGC GCATGCACCT

    NotI
    ~~~~~~
    . P V A A A E P L K D E L *
911 GGCCGGTCGC GGCCGCGGAA CCACTGAAGG ATGAGCTGTA AAGAAGCAGA TCGTTCAAAC ATTTGGCAAT
981 AAAGTTTCTT AAGATTGAAT CCTGTTGCCG GTCTTGCGAT GATTATCATA TAATTTCTGT TGAATTACGT
1051 TAAGCATGTA ATAATTACAA TGTAAATGCAT GACGTATATT ATGAGATGGG TTTTATATGAT TAGAGTCCCG
1121 CAATTATACA TTTAATACGC GATAGAAAAC AAAATATAGC GCGCAAACTA GGATAAATTA TCGCGCGCGG

    HindIII
    ~~~~~~
    ClaI          XbaI
    ~~~~~~
1191 TGTCACTCTAT GTTACTAGAT CGATAAGCTT CTAGAGCGGC CGGTGGAGCT CCAATTCGCC CTATAGTGAG
1261 TCGTATTACG CGGCTCACT GGCCTGTCGTT TTACAAACGTC GTGACTGGGA AAACCTTGGC GTTACCCAAAC
1331 TTAATCGCCT TGCAGCACAT CCCCCTTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCGCA CCGATCGCCC
1401 TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGACGCG CCCTGTAGCG CGCATTAAG CGCGCGCGGT
1471 GTGGTGGTTA CGCGCAGCGT GACCGCTACA CTTGCCAGCG CCCTAGCGCC CGCTCCCTTC TAGGTTCTTC
1541 CTTCCTTTCT CGCACGTTT TACGGGCACC TCGACCCCAA AAACTTGAT TAGGGTGATG GTTACAGTAG TGGGCCATCG
1611 ATTTAGTGCT TTACGGCACC CGGTTTTTCG CCCTTTGACG TTGGAGTCCA CGTTCCTTAA TAGTGGACTC TTGTTCCAAA
1681 CCCTGATAGA CGGTTTTCG ACTCAACCTT ATCTCGGTCT ATTCCTTTGA TTTATAAGGG ATTTGCCGA TTTCGGCCTA
1751 CTGGAACAAC ACTCAACCTT ATCTCGGTCT ATTCCTTTGA TTTATAAGGG ATTTGCCGA TTTCGGCCTA
1821 TTGGTTAAA AATGAGCTGA TTTAACAAA ATTTAACGCG AATTTTAAAC AAATATTAAC GCTTACAAT
1891 TAGGTGGCAC TTTTCGGGA AATGTGCGG GAACCCCTAT TTGTTTATTT TTCTAAATAC ATTCAAATAT
1961 GTATCCGCTC ATGAGACAAT AACCTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT
2031 CAACATTTC GTGTCGCCCT TATTCCCTTT TTTGCGGCAT TTTGCTTCC ACGAGTGGGT TACATCGAAC TGGATCTCAA
2101 CGCTGGTGAA AGTAAAGAT GCTGAAGATC AGTTGGGTGC AGTGGGTGC TTTCCCAATGA TGAGCACTTT TAAAGTTCTG
2171 CAGCGGTAAG ATCCTTGAGA GTTTTCGCC CGAAGAACGT TTTCCCAATGA TGAGCACTTT TAAAGTTCTG
2241 CTATGTGGCG CGGTATTATC CCGTATTGAC GCCGGGCAAG AGCAACTCGG TCGCCGCATA CACTATTTCT
2311 AGAATGACTT GGTGAGTAC TCACCAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG TAAGAGAAAT
2381 ATGCAGTGCT GCCATAACCA TGAGTGATTA CACTGCGGCC AACTTACTTC TGACAACGAT CGGAGGACCG
```

FIG._34C

48 / 154

```
2451 AAGGAGCTAA CCGCTTTTTT GCACAACATG GGGATCATG TAACTCGCCT TGATCGTTGG GAACCGGAGC
2521 TGAATGAAGC CATACCAAAC GACGAGCGTG ACACCACGAT GCCTGTAGCA ATGGCAACAA CGTTGCGCAA
2591 ACTATTAACT GGCGAACTAC TTACTCTAGC TTCCCAGCAA CAATTAATAG ACTGGATGGA GCGCGATAAA
2661 GTTGACAGGAC CACTTCTGCG CTCGGCCCTT CCGGCTGGCT GGTATTATTGC TGATAAATCT GGAGCCGGTG
2731 AGCGTGGGTC TCGCGGTATC ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC TCCCGTATCG TAGTTATCTA
2801 CACGACGGGG AGTCAGGCAA CTATGGATGA ACGAATAGA CAGATCGCTG AGATAGGTGC CTCACCTGATT
2871 AAGCATTTGT AACTGTCAGA CCAAGTTTAC TCATATATAC TTTAGATTGA TTTAAAACTT CATTTTAAAT
2941 TTAAGAAGGAT CTAGGTGAAG ATCCTTTTGT TAGAAAAGAT CAAAGGATCT GACCAAAATC CTTAACGTG
3011 CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT CAAAGGATCT TCTTGAGATC CTTTTCGTT
3081 TGCTGCTTGC AAACAAAAAA ACCACCGCTA CCAGCGTGG TTTGTTTGCC GGATCAAGAG CTACCAACTC
3151 TTTTTCGAA GGTAACTGGC TTCAGCAGAG CGCAGATACC AAATACTGTC CTTCTAGTGT AGCCGTAGTT
3221 AGGCCACCAC TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCCCTGTT ACCAGTGGCT
3291 GCTGCCAGTG CGGATAAGTC GTGTCTTACC GGGTTGGACT CAAGACGATA GTTACCAGAT AAGCGCAGC
3361 GGTGCGGCTG AACGGGGGT TCCTGTCACAC AGCCAGCTT GGAGCGAAGG ACCTACACCG AACTGAGATA
3431 CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA TCCGGTAAAGC
3501 GGCAGGGTCG GAACAGGAGA GCGCACGAGG GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCCCTG
3571 TCGGGTTTCG CCACCTCTGA CTTGAGCGTC GATTTTGTG ATGCTCGTCA GGGGGCGGA GCCTATGGAA
3641 AAACGCCAGC AACCGGCCCT TTTTACGGTT CCTGGCCCTT TGCTGACCAT TTGCTCTTCTT
3711 GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATACCGCT CGCCGACGCC
3781 GAACGACCGA GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA AGAGCGCCA ATACGCAAAAC CGCCTCTCCC
3851 CGCGCGTTGG CCGATTCAAT AATGCAGCTG GCACGACAGG TTTCCCGACT GGAAGCGGG CAGTGAGCGC
3921 AACGCAATTA ATGTGAGTTA GCTCACTCAT TAGGCACCCC AGGCTTTACA CTTTATGCTT CCGGCTCGTA
3991 TGTTGTGTGG AATTGTGAGC GGATAACAAT TTCACACAGG AAACAGCTAT GACCATGATT ACGCCAAGCG

                                KpnI
                                ~~~~~
4061 CGCAATTAACT CCTCACTAAA GGAACAACAA GGTGGGTACC GGGCCCCCCC TCAGAGTCAT TCATATGCTT
4131 GAGAAGAGAG TCGGGATAGT CCAAAATAAA ACAAAAGTAA GATTACCTGG TCAAAAGTGA AAACATCAGT
4201 TAAAAGGTGG TATAAGTAAA ATATCGGTAA TAAAAGGTGG CCCAAAGTGA AATTACTCT TTTCTACTAT
4271 TATAAAAAAT GAGGATGTTT TGTCGGTACT TTGATACGTC ATTTTGTGAT GAATTGGTTT TTAAGTTTAT
4341 TCGCGATTGG GAAATGCATA TCTGTATTGG AGTCGGTTT TAAGTTCGTT GCTTTGTAA ATACAGAGGG
4411 ATTTGTATAA GAAATATCTT TAAAAAACCC ATATGCTAAT TTGACATAAT TTTTGAGAAA AATATATATT

                                EcoRI
                                ~~~~~
4481 CAGGCGAATT CCACAATGAA CAATAATAAG ATTAAATAG CTTGCCCCCG TTGCAGCGAT GGGTATTTTT
4551 TCTAGTAAAA TAAAAGATAA ACTTAGACTC AAAACATTTA CAAAACAAAC CCTAAAGTC CTAAAGCCCC
```

FIG.-34D


```
4621 AAGTGCTATG CACGATCCAT AGCAAGCCCA GCCCAACCCA ACCCAACCCA ACCCAGCCCA GTGCAGCCAA
4691 CTGGCAAATA GTCTCCACCC CCGGCACTAT CACCGTGAGT TGTCCGCACC ACCGCACGTC TCGCAGCCAA
4761 AAAAAAAAAA AGAAGAAAA AAAAGAAAA GAAAGACAGC AGGTGGGTCC GGTCTGTGGG GGCCGGAAAA
4831 GCGAGGAGGA TCGCGAGCAG CGACGAGGCC CGGCCCTCCC TCCGCTTCCA AAGAAACGCC CCCCATCGCC
4901 ACTATATACA TACCCCTCCC TCTCTCTCCA TCCCCCACA CCTACCACCA CCACCACCCAC CACCTCCTCC
4971 CCCCTCGCTG CCGGACGACG AGCTCCTCCC CCCTCCCCCT CGCCGCGCCG CGGTAACCCAC CCCGCCCCTC
5041 TCCCTCTTCT TTCTCCGTTT TTTTCTTCGT CTCGGTCTCG ATCTTTGGCC TTGGTAGTTT GGGTGGGCGA
5111 GAGCGGCTTC GTCGCCCAGA TCGGTGCGCG GGAGGGGCGG GATCTCGCGG CTGGCGTCTC CGGGCGTGAG

          BamHI          BglII
          ~~~~~          ~~~~~
5181 TCGGCCCCGA TCCTCGCGGG GAATGGGGCT CTCGGATGTA GATCTTCTTT CTTTCTTCTT TTTGTGGTAG
5251 AATTGAATC CCTCAGCATT GTTCATCGGT AGTTTTTCTT TTCATGATTT GTGACAAATG CAGCCCTCGTG
5321 CGGAGCTTTT TTGTAGC
```

FIG._34E

50 / 154

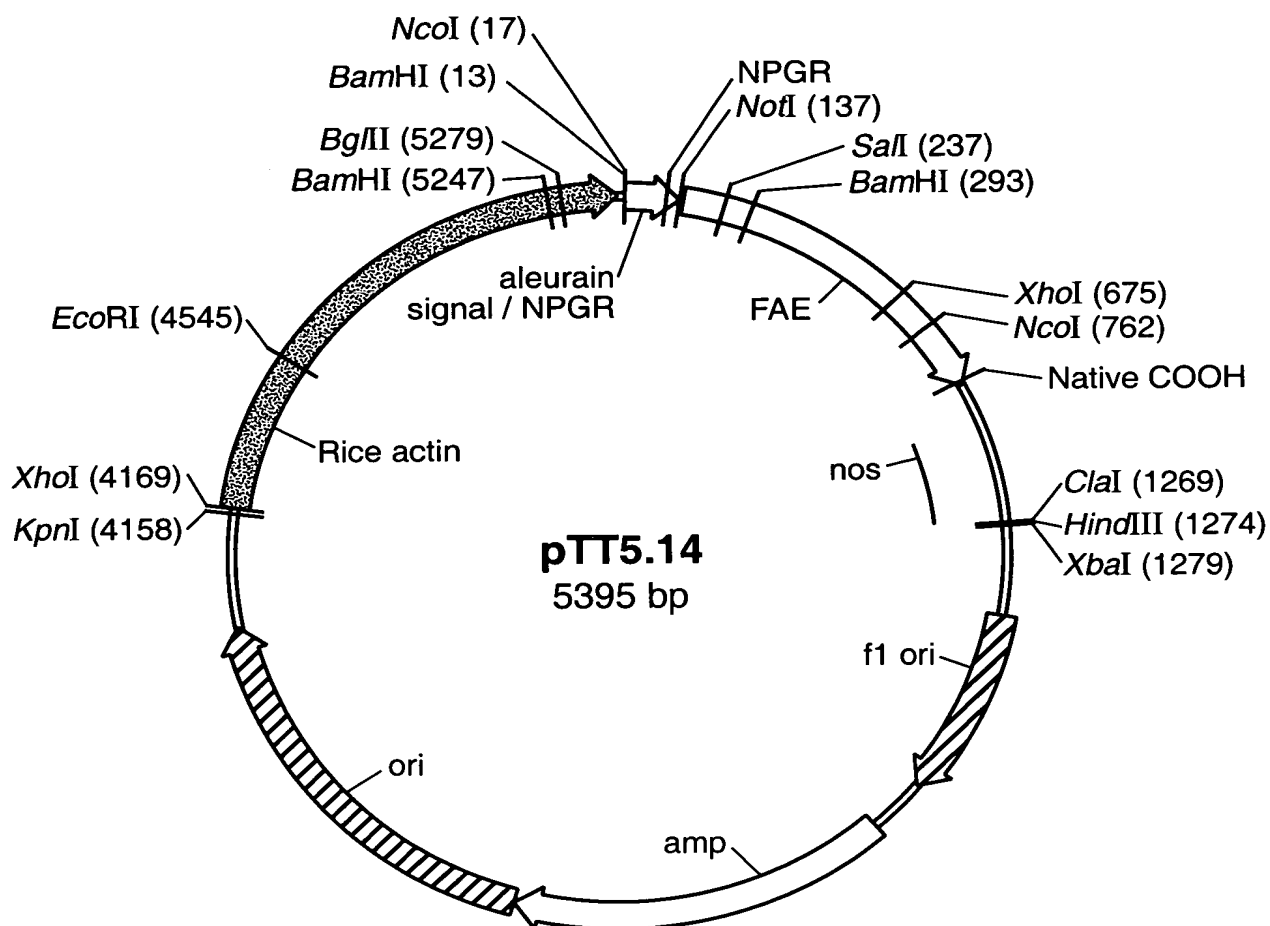


FIG._35A

```

                                NcoI
                                ~~~~~
                                BamHI
                                ~~~~~
1  CCTGACGCCG AGGATCCATG GCCACAGCCC GCGTCCTCCT CCTLGCGCTC GCCGTGCTGG CCACGGCCGC
   M A H A R V L L L A L A V L A T A A .
   NotI
71 CGTCGCCGTC GCCTCCTCCT CCTCCTTCGC CGACTCCAAC CCGGCCGGC CCGTCACCGA CCGCGCGGCC
   . V A V A S S S F A D S N P G R P V T D R A A
   NotI
   ..
141 GCCTCCACGC AGGCATCTC CGAAGACCTC TACAGCCGTT TAGTCGAAAT GGCCACTATC TCCCAAGCTG
   A S T Q G I S E D L Y S R L V E M A T I S Q A A
   Sali
   ~~~~~
211 CCTACGCCGA CCTGTGCAAC ATTCCGTCGA CTATTATCAA GGGAGAGAAA ATTACAAATT CTCAAACTGA
   . Y A D L C N I P S T I I K G E K I Y N S Q T D .
   BamHI
   ~~~~~
281 CATTAACGGA TGGATCCTCC GCGACGACAG CAGCAAGAA ATAATCACCG TCTTCCGTGG CACTGGTAGT
   D T N L Q L D T N Y T L T P F D T L P Q C N G C
351 GATACGAATC TACAACCTCGA TACTAACTAC ACCCTCAGC CTTTCGACAC CCTACCACAA TGCAACGGTT
   . E V H G G Y I G W V S V Q D Q V E S L V K Q .
421 GTGAAGTACA CCGTGGATAT TATATTGGAT GGGTCTCCGT CCAGGACCAA GTCGAGTCGC TTGTCAAACA
   . Q V S Q Y P D Y A L T V T G H X L G A S L A A
491 GCAGGTTAGC CAGTATCCGG ACTACGCGCT GACCGTGACC GGCCACKCC TCGGCGCCTC CCTGGCGGCA
   L T A A Q L S A T Y D N I R L Y T F G E P R S G
561 CTCACCTGCC CCGAGCTGTC TCGACATAC GACAACATCC GCCTGTACAC CTTCCGGCGAA CCGCGCAGCG
   XhoI
   ~~~~~
631 GCAATCAGGC CTTGCGCTCG TACATGAACG ATGCCTTCCA AGCCTCGAGC CCAGATACGA CGCAGTATTT
   . N Q A F A S Y M N D A F Q A S S P D T T Q Y F .
   NotI
   ~~~~~
   . R V T H A N D G I P N L P P V E Q G Y A H G G

```

FIG._35B

52 / 154

```

701 CCGGGTCACT CATGCCAACG ACGGCATCCC AAACCTGCCC CCGGTGGAGC AGGGGTACGC CCATGGCGGT
    V E Y W S V D P Y S A Q N T F V C T G D E V Q C
771 GTAGAGTACT GGAGCGTTGA TCCTTACAGC GCCAGAACCA CATTGTCTTG CACTGGGGAT GAAGTGCAGT
    . C E A Q G G Q G V N A H T T Y F G M T S G A .
841 GCTGTGAGGC CCAGGGCGGA CAGGGTGTGA ATAATGCGCA CACGACTTAT TTTGGGATGA CGAGCGGAGC
    . C T W *
911 CTGTACATGG TGATCAGTCA TTTCAGCCTC CCCGAGTGTA CCAGGAAAGA TGGATGTCCT GGAGAGGGGG
981 CCGCGTAACC ACTGAAGGAT GAGCTGTAAA GAAGCAGATC GTTCAAAACAT TTGGCAATAA AGTTTCTTAA
1051 GATTGAATCC TGTGCGCGT CTGCGATGA TTATCATATA ATTTCTGTTG AATTACGTTA AGCATGTAAT
1121 AATTAAACATG TAATGCATGA CGTTATTAT GAGATGGGTT TTATGATTA GAGTCCCGCA ATTATACATT
1191 TAATACGCGA TAGAAAACAA AATATAGCGC GCAAACTAGG ATAAATTATC GCGCGCGGTG TCATCTATGT
    HindIII
    ~~~~~

```

Clal XbaI

```

1261 TACTAGATCG ATAAGCTTCT AGAGCGGGCG GTGGAGCTCC AATTCGCCCT ATAGTGAGTC GTATTACGG
1331 CGCTCACTGG CCGTCGTTTT ACAACGTCGT GACTGGGAAA ACCCTGGCGT TACCCAACTT AATCGCCTTG
1401 CAGCACATCC CCTTTTCGCC AGCTGGCGTA ATAGCGAAGA GCGCCGCACC GATCGCCCTT CCCAACAGTT
1471 GCGCAGCCTG AATGGCGAAT GGGACGCGCC CTGTAGCGGC GCATTAAGCG CGCGGGGTGT GGTGGTTACG
1541 CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCC GTCCCTTTCG CTTCTTCCCT TCCTTTCTCG
1611 CCACGTTTCG CCGCTTTCCT CGTCAAGCTC TAAATCGGG GCTCCCTTTA GGGTTCGGAT TTAGTGCTTT
1681 ACGGCACCTC GACCCCAAAA AACTTGATTA GGGTATGGT TCACGTAGTG GGCCATCGCC CTGATAGACG
1751 GTTTTTCGCC CTTTGACGTT GGAGTCCACG TTCTTTAATA GTGGACTCTT GTTCCAACCT GGAACAACAC
1821 TCAACCCCTAT CTCGGTCTAT TCTTTTGATT TATAAGGGAT TTTGCCGAT TCGGCCCTAT GGTAAACAAA
1891 TGAGCTGATT TAACAAAAAT TTAACGCGAA TTTTAAACAA ATATTAAACG TTACAATTTA GTGGCACATT
1961 TTCGGGAAA TGTGCGCGGA ACCCTATT TTTTATTTT CTAAATATGT TCAAAATATG ATCCGCTCAT
2031 GAGACAATAA CCTGATAAA TGCTTCAATA ATATTGAAA AGGAAGAGTA TGAGTATTCA ACATTTCCGT
2101 GTCGCCCTTA TTCCCTTTT TCGGCGCATTT TGCCCTTCCTG TTTTTCCTCA CCCAGAAAAG CTGGTGAAG
2171 TAAAAGATGC TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAACTG GATCTCAACA GCGGTAAGAT
2241 CCTTGAGAGT TTTCGCCCCG AAGAACGTTT TCCAATGATG AGCACTTTTA AAGTCTGCT ATGTGGCGCG
2311 GTATTATCCC GTATTGACGC CGGGCAAGAG CAACTCGGTC GCCGCATACA CTATTCTCAG AATGACTTGG
2381 TTGAGTACTC ACCAGTCACA GAAAAGCATC TTACGGATGG CATGACAGTA AGAGAAATTAT GCAGTGCTGC
2451 CATAACCATG AGTGATAACA CTGCGGCCAA CTTACTTCTG ACAACGATCG GAGGACCGAA GGAGCTAACC
2521 GCTTTTTTGC ACAACATGGG GGATCATGTA ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA
2591 TACCAAAACGA CGAGCGTGAC ACCACGATGC CTGTAGCAAT GGCAACAACG TTGCGCAAC TATTAAGTGG
2661 CGAACTACTT ACTCTAGCTT CCCGGCAACA ATTAATAGAC TGGATGGAGG CGGATAAAGT TGCAGGACCA

```

FIG._35C

```

2731 CTTCTGCGCT CGGCCCTTCC GGCTGGCTGG TTTATTGCTG ATAAATCTGG AGCCGGTGAG CGTGGGTCTC
2801 GCGGTATCAT TGCAGCACTG GGGCCAGATG GTAAGCCCTC CCGTATCGTA GTTATCTACA CGACGGGGAG
2871 TCAGGCAACT ATGGATGAAC GAAATAGACA GATCGCTGAG ATAGGTGCCCT CACTGATTAA GCATTGGTAA
2941 CTGTCAGACC AAGTTTACTC ATATATACTT TAGATTGATT TAAAACTTCA TTTTAAATTT AAAAGGATCT
3011 AGGTGAAGAT CCTTTTGTAT AATCTCATGA CCAAAATCCC TTAACGTGAG TTTTCGTTCC ACTGAGCGTC
3081 AGACCCCGTA GAAAAGATCA AAGGATCTTC TTGAGATCCT TTGAGATCTG GCGTAATCTG CTGCTTGCAA
3151 ACAAATAAAC CACCGCTACC AGCGGTGGTT TGTGTCGGG ATCAAGAGCT ACCAACTCTT TTTCGGAAGG
3221 TAACCTGGCT CAGCAGAGCG CAGATACCAA ATACTGTCTT CGCTCTGCTA ATCCTGTTAC CAGTGGCTGC TGCCAGTGGC
3291 CAAGAACTCT GTAGCACCCG GTTGACTCA AGACGATAGT TACCGGATAA GCGGCAGCGG CTGAGATACC TACAGCGTGA
3361 GATAAGTCGT GTCTTACCGG GTTGACTCA AGACGATAGT TACCGGATAA GCGGCAGCGG CTGAGATACC TACAGCGTGA
3431 CGGGGGTTT GTGCACACAG CCCAGCTTGG AGCGAAGCG GAGAGGATC GGTATCTTTA TAGTCCCTGC GGGTTTCGCC
3501 GCTATGAGAA AGCGCCACGC GCACGAGGGA GCTTCCAGGG GGTATCTTTA TAGTCCCTGC GGTATCTGCC GGTATCTGCC
3571 ACAGGAGAGC GCACGAGGGA GCTTCCAGGG GGTATCTTTA TAGTCCCTGC GGTATCTGCC GGTATCTGCC
3641 ACCTCTGACT TGAGCGTCCA TTTTGTGTAT GCTCGTCAGG GGTATCTTTA TAGTCCCTGC GGTATCTGCC GGTATCTGCC
3711 CGCGGCTTTT TTACGGTTCC TGCGCTTTTG CTGGCTTTT GCTCACATGT TCTTTCCTGC GGTATCTGCC GGTATCTGCC
3781 GATTCGTGG ATAAACCGTAT TACCGCTTTT GAGTGAGCTG ATACCGCTCG CCGCAGCCGA ACGACCGAGC
3851 GCAGCGAGTC AGTGAGCGAG GAAGCGGAAG AGCGCCCAAT ACGCAAAACG CCGTCTCCCG CCGCTTGGCC
3921 GATTCATTAA TGCAGCTGGC ACGACAGGTT FCCCGACTGG AAAGCGGCA GTGAGCGCAA CGCAATTAAT
3991 GTGAGTTAGC TCACTCATTA GGCACCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA
4061 TTGTGAGCGG ATAACAATTT CACACAGGAA ACAGCTATGA CCATGATTAC GCCAAGCGCG CAATTAACCC

                                KpnI                                XhoI
                                ~~~~~                                ~~~~~
4131 TCACTAAAGG GAACAAAAGC TGGGTACCGG GCGCCCTCTC GAGGTCAATC ATATGCTTGA GAAGAGAGTC
4201 GGGATAGTCC AAAATAAAAC AAAGGTAAGA TTACCTGGTC AAAAGTGAAA ACATCAGTTA AAAGGTGGTA
4271 TAAGTAAAT ATCGGTAATA AAAGGTGGCC CAAAGTGAAA TTTACTCTTT TCTACTATTA TAAAAAATGA
4341 GGATGTTTTG TCGGTACTTT GATACGTCAAT TTTTGTATGA ATTGGTTT AAGTTTATTC GCGATTGGGA
4411 AATGCATATC TGTATTTGAG TCGGTTTTTA AGTTCGTTGC TTTTGTAAAT ACAGAGGGAT TTGTATAAGA

                                EcorI
                                ~~~~~
4481 AATATCTTTA AAAAACCCAT ATGCTAATTT GACATAATTT TTGAGAAAAA TATATATTC GCGCAATTCC
4551 ACAATGAACA ATAATAAGAT TAAAATAGCT TGCCCTCCGTT GCAGCGATGG GTATTTTTTC TAGTAAAAA
4621 AAAGATAAAC TTAGACTCAA AACATTTACA AAAACAACCC CTAAGTCCCT AAAGCCCAA GTGCTATGCA
4691 CGATCCATAG CAAGCCGAGC CCAACCCCAAC CCAACCCCACT GCAGCCCACT GGCAATAAGT
4761 CTCCACCCCC GGCATCATCA CCGTGAGTTG TCCGCACCCAC CGCACGCTC GCAGCCCAA AAAAATAAG
4831 AAAGAAAAAA AAGAAAAAGA AAAACAGCAG GTGGGTCCCG GTGCTGGGG CCGGAAAAAGC GAGGAGGATC
    
```

FIG._35D

4901 GCGAGCAGCG ACGAGGCCCG GCCCTCCCTC CGCTTCCAA GAAACGCCCC CCATCGCCAC TATATACATA
4971 CCCCCCCTC TCCTCCCATC CCCCACCC TACCACACC ACCACCACCA CCTCCTCCCC CCTCGCTGCC
5041 GGACGACGAG CTCTCCCTCC CTCCCTCCCTC GCGCGGCCG GTAACCAACC CGCCCCCTCTC CTCCTTCTTT
5111 CTCCGTTTTT TTTTTCGTCT CGGTCTCGAT CTTTGGCCTT GGTAGTTGG GTGGGCGAGA GCGGCTTCGT
BamHI
5181 CGCCACAGATC GGTGCGCGGG AGGGCGGGA TCTCGCGGCT GCGTCTCTCCG GCGGTGAGTC GGCCCGGATC
BamHI
~
5251 CTCGCGGGA ATGGGCTCT CGGATGTAGA TCTTCTTTCT TTCTTCTTTT TGTGGTAGAA TTTGAATCCC
5321 TCAGCATTTG TCATCGGTAG TTTTCTTTT CATGATTGT GACAAATGCA GCCTCGTGCG GAGCTTTTTT
5391 GTAGC

FIG._35E

55 / 154

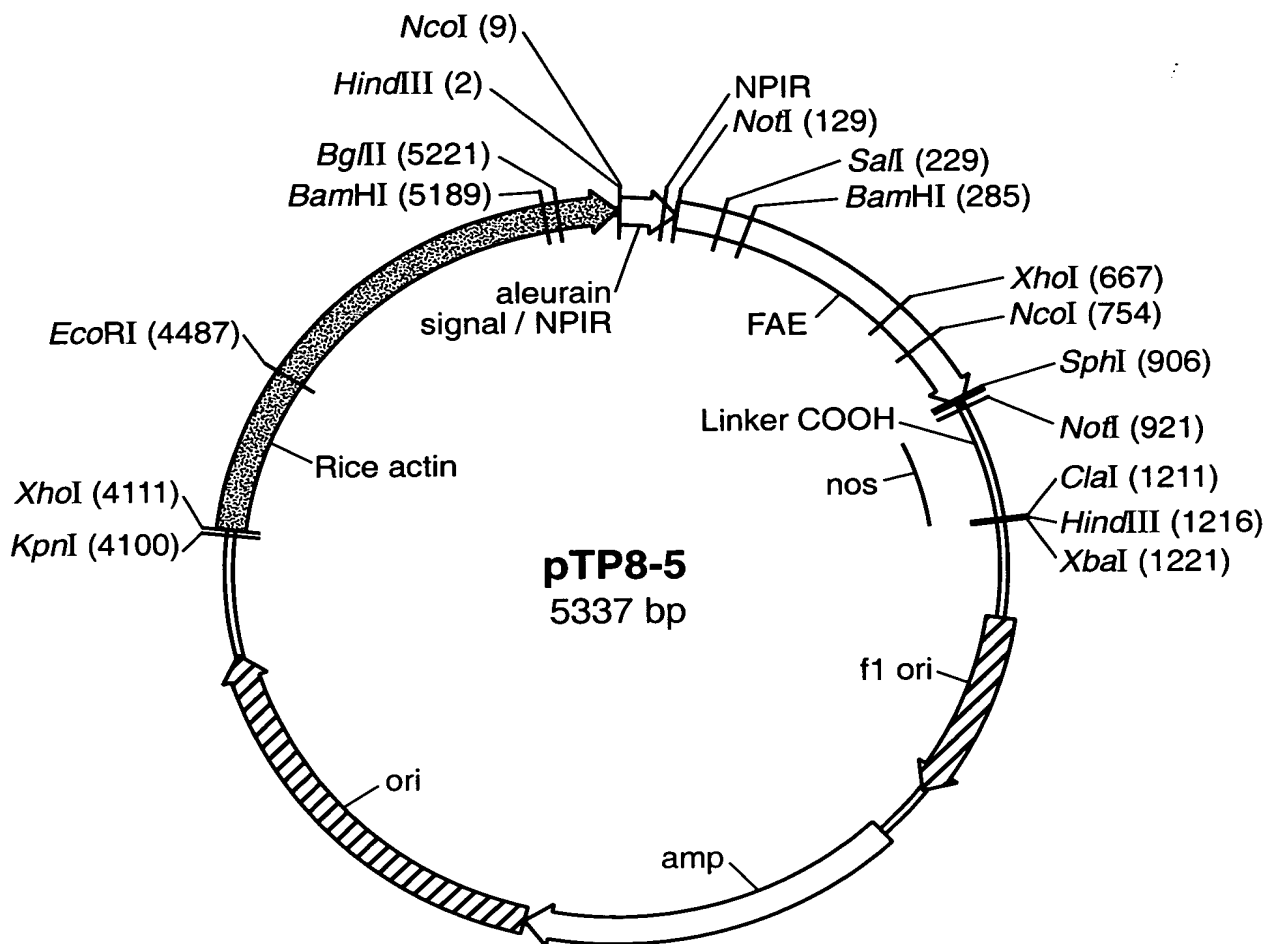


FIG._36A

FIG. 36B

57 / 154

```

701 CTCATGCCAA CGACGGCATC CCAAACCTGC CCCCGGTGGA GCAGGGGTAC GCCCATGGCG GTGTAGAGTA
    . W S V D P Y S A Q N T F V C T G D E V Q C C E
771 CTGGAGCGTT GATCCTTACA GCGCCAGAA CACATTGTC TGCACCTGGG ATGAAGTGCA GTGCTGTGAG
    SphI
    ~~~~
    A Q G G Q G V N N A H T T Y F G M T S G A C T W
841 GCCCAGGGCG GACAGGGTGT GAATAATGCG CACACGACTT ATTTGGGAT GACGAGCGGC GCATGCACCT
    NotI
    ~~~~~~
    . P V A A A *
911 GGCCGGTCGC GGCCGCGTAA CCACTGAAGG ATGAGCTGTA AAGAAAGCAGA TCGTTCAAAC ATTTGGCAAT
981 AAAGTTCTTT AAGATTGAAT CCTGTGCGG GTCTTGCGAT GATTATCATA TAAATTCGTG TGAATTACGT
1051 TAAGCATGTA ATAATTAAACA TGTAAATGCAT GACGTTATTT ATGAGATGGG TTTTATATGAT TAGAGTCCCG
1121 CAATTATACA TTTAATACGC GATAGAAAAC AAAATATAGC GCGCAAACTA GGATAAATTA TCGCGCGCGG
    HindIII
    ~~~~~~
    ClaI XbaI
    ~~~~~~
1191 TGTCACTCTAT GTTACTAGAT CGATAAGCTT CTAGAGCGGC CGGTGGAGCT CCAATTCGCC CTATAGTGAG
1261 TCGTATTACG CGGCTCACT GGCGTCTGTT TTACAACGTC GTGACTGGA AAACCTTGGC GTTACCCAAAC
1331 TTAATCGCCT TGCAGCACAT CCCCTTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC
1401 TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGGACGCG CCCTGTAGCG GCGCATTAAG CGCGGCGGGT
1471 GTGGTGGTTA CGCGCAGCGT GACCGCTACA CTTGCCAGCG CCCTAGCGCC CGCTCCCTTC GCTTTCCTCC
1541 CTTCCCTTCT TACGGCACC TCGACCCCAA AAACTTGAT TAGGGTGATG GTTCACGTAG TGGGCCATCG
1611 ATTTAGTGCT TTACGGCACC CCCTTTTCG ATCTCGGTCT TTGGAGTCCA CGTTCCTTAA TAGTTCCTAA
1681 CCCTGATAGA CGGTTTTCG CCCTTTTCG ATCTCGGTCT TTGGAGTCCA CGTTCCTTAA TAGTTCCTAA
1751 CTGGAACAAC ACTCAACCTT AATGAGCTGA TTTAACAAA AATTAACGCG AATTTTAAAC AAATATAAT
1821 TTGGTTAAA AATGAGCTGA TTTTCGGGA AATGTCGCGG GAACCCCTAT TTGCTTAAAC AATTAATAT
1891 TAGGTGGCAC TTTTCGGGA AATGTCGCGG GAACCCCTAT TTGCTTAAAC AATTAATAT
1961 GTATCCGCTC ATGAGACAAT AACCTTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTAT
2031 CAACATTTCC GTGTCGCCCT TATTCCCTTT TTTGCGGCAT TTGCTTAAAC AATTAATAT
2101 CGCTGGTGAA AGTAAAGAT GCTGAAGATC AGTTGGGTGC TTTGCTTAAAC AATTAATAT
2171 CAGCGGTAG ATCCTTGAGA GTTTTCGCC CGAAGAACGT TTTCCAAATGA TGAGCACTTT TAAAGTTCTG
2241 CTATGTGGCG CGGTATTATC CCGTATTGAC GCCGGGCAAG AGCAACTCGG TCGCCGCATA CACTATTCTC
2311 AGAATGACTT GGTGAGTAC TCACCAGTCA CAGAAAAGCA TCTTACGGAT GGCAATGACAG TAAGAGAAAT

```

FIG._36C

58 / 154

```

2381 ATGCAGTGCT GCCATAACCA TGAGTGATAA CACTGCGGCC AACTTACTTC TGACAAACGAT CGGAGGACCG
2451 AAGGAGCTAA CCGCTTTTTC GCACAAACATG GGGGATCATG TAACTCGCCT TGATCGTTGG GAACCGGAGC
2521 TGAATGAAGC CATACCAAAC GACGAGCGTG ACACCACGAT GCCTGTAGCA ATGGCAACAA CGTTGCGCAA
2591 ACTATTAACT GCGAAACTAC TTACTCTAGC TTCCCGGCAA CAATTAATAG ACTGGATGGA GCGGATATAA
2661 GTTGCAGGAC CACTTCTGCG CTCGGCCCTT CCGGCTGGCT TGGTTATTGC TGATAAATCT GGAGCCGGTG
2731 AGCGTGGGTC TCGCGGTATC ATTGCAGCAC ATATGGATGA TGGTAAGCCC TCCCGTATCG TAGTTATCTA
2801 CACGACGGGG AGTCAGGCAA CTAAGTTTAC TCATATATAC CAGATCGCTG AGATAGGTGC CTCACCTGATT
2871 AAGCATTTGGT AACTGTGAGA CCAAGTTTAC CCAATCTCAT TTTAGATTGA TTTAAAACTT CATTTTAAAT
2941 TTAAGAGGAT CTAAGTGAAG ATCCTTTTTC TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTAAACGT AGTTTTCGTT
3011 CCACTGAGCG TCAGACCCCG TCAGACCCCG ACCACCGCTA CCAGCGGTGG TTTGTTTGCC GGATCAAGAG CTACCAACTC
3081 TGCTGCTTGC AAACAAAATA GGTAACTGGC TTCAGCAGAG CGCAGATACC AAATACTGTC TTTGTTTGCC GGATCAAGAG CTACCAACTC
3151 TTTTCCGAA AGCCACACAC TTCAAGAACT CTGTAGCACC GGTGCTTACC GGTGCTTACC CTTGTTTGCC GGATCAAGAG CTACCAACTC
3221 AGGCCACAC GCTGCCAGT AACGGGGGT TCCTGACAC CAGCTTCCAG GAGCTTCCAG GGTGCTTACC CTTGTTTGCC GGATCAAGAG CTACCAACTC
3291 GCTGCCAGT AACGGGGGT TCCTGACAC CAGCTTCCAG GAGCTTCCAG GGTGCTTACC CTTGTTTGCC GGATCAAGAG CTACCAACTC
3361 GGTGCCAGT AACGGGGGT TCCTGACAC CAGCTTCCAG GAGCTTCCAG GGTGCTTACC CTTGTTTGCC GGATCAAGAG CTACCAACTC
3431 CTAACAGCT GAGCTATGAG AAAGCGCAC CAGCTTCCAG GAGCTTCCAG GGTGCTTACC CTTGTTTGCC GGATCAAGAG CTACCAACTC
3501 GGCAGGCTG GAGCTATGAG AAAGCGCAC CAGCTTCCAG GAGCTTCCAG GGTGCTTACC CTTGTTTGCC GGATCAAGAG CTACCAACTC
3571 TCGGGTTTCG CCACCTCTGA CTTGAGCGTC TTTTACGGTT CTTGAGCGTC TTTGAGCGTC TTTGAGCGTC TTTGAGCGTC TTTGAGCGTC
3641 AAACGCCAGC AACGGGGGT TTTTACGGTT CTTGAGCGTC TTTTACGGTT CTTGAGCGTC TTTGAGCGTC TTTGAGCGTC TTTGAGCGTC
3711 GCCTTATCCC CAGCTTCTGT GGTAAACCGT TCAGTGAGCG AGGAAGCGGA AGAGCGGCGA AGAGCGGCGA AGAGCGGCGA AGAGCGGCGA
3781 GAACGACCGA GCGCAGCGAG CCGATTCATT AATGCAGCTG GCACGACAGG TTTCCCGGCT TTTCCCGGCT TTTCCCGGCT TTTCCCGGCT
3851 CGCGCGTTGG CCGATTCATT AATGCAGCTG GCACGACAGG TTTCCCGGCT TTTCCCGGCT TTTCCCGGCT TTTCCCGGCT TTTCCCGGCT
3921 AACGCAATTA ATGTGAGTTA GCTCACTCAT TAGGCACCCC AGGCTTTACA CTTTATGCTT GACCATGATT ACGCCAAGCG
3991 TGTTGTGTGG AATTGTGAGC GGATAACAAT TTCACACAGG AACACAGCTAT GACCATGATT ACGCCAAGCG

4061 CGCAATTAA CTTCACTAAA GGGAAACAAA GCTGGGTACC GGGCCCCCCC TCGAGGTGAT TCATATGCTT
4131 GAGAAGAGAG TCGGGATAGT CCATAATAAA ACAAAAGTAA GATTACCTGG TCAAAAAGTGA AAACATCAGT
4201 TAAAAGGTGG TATAAGTAAA ATATCGGTAA TAAAAGGTGG CCATAAGTGA AATTACTCTT TTTCTACTAT
4271 TATAAAAAAT GAGGATGTTT TGTCGGTACT TTGATACGTC ATTTTGTGAT GAATTGGTTT TTAAGTTTAT
4341 TCGCGATTG GAAATGCATA TCTGTATTG AGTCGGTTT TAAAGTTCGTT GCTTTGTGAA ATACAGAGGG
4411 ATTTGTATTA GAAATATCTT TAAAAAACCC ATATGCTAAT TTGACATAAT TTTTGAGAAA AATATATATT

      EcoRI
      ~~~~~
4481 CAGGCGAATT CCACAATGAA CAATAATAAG ATTAAAAATAG CTTGCCCCCG TTGCAGCGAT GGGTATTTTT

```

XhoI

KpnI

FIG._36D

```
4551 TCTAGTAAAA TAAAAGATAA ACTTAGACTC AAAACATTTA CAAAAACAAC CCTAAAGTC CTAAAGCCCA
4621 AAGTGCTATG CACGATCCAT AGCAAGCCCA GCCCAACCCA ACCCAACCCA ACCCAACCCA GTGCAGCCAA
4691 CTGGCAAATA GTCTCCACCC CCGGCACTAT CACCGTGAGT TGTCCGCACC ACCGCACGTC TCGCAGCCAA
4761 AAAAAAATA AGAAAGAAAA AAAGAAAAA AAAAGAAAA CGGCGGTCC GGGTCGTGGG GGCCGGAAAA
4831 GCGAGGAGGA TCGCGAGCAG CGACGAGGCC CGGCCCTCCC TCCGCTTCCA AAGAAACGCC CCCCATCGCC
4901 ACTATATACA TACCCCCCCC TCTCCTCCCA TCCCCCAAC CTAACCAACA CCACCACCAC CACCTCCTCC
4971 CCCCTCGCTG CCGGACGACG AGCTCCTCCC CCCTCCCCCT CCGCCGCCGC CGGTAACCAAC CCCGCCCTC
5041 TCCTCTTTCT TTCTCCGTTT TTTTTCGTCG CTCGGTCTCG ATCTTTGGCC TTGGTAGTTT GGGTGGGCGA
5111 GAGCGGCTTC GTCGCCCAGA TCGGTGCGCG GGAGGGGCGG GATCTCGCGG CTGGCGTCTC CGGGCGTGAG

          BamHI          BglII
          ~~~~~          ~~~~~

5181 TCGGCCCCGA TCCTCGCGGG GAATGGGGCT CTCGGATGTA GATCTTCTTT CTTCCTTCTT TTTGTGGTAG
5251 AATTGAATC CCTCAGCAAT GTTCATCGGT AGTTTTTCTT TTCATGATTT GTGACAAATG CAGCCTCGTG
5321 CGGAGCTTTT TTGTAGC
```

FIG. 36E

60 / 154

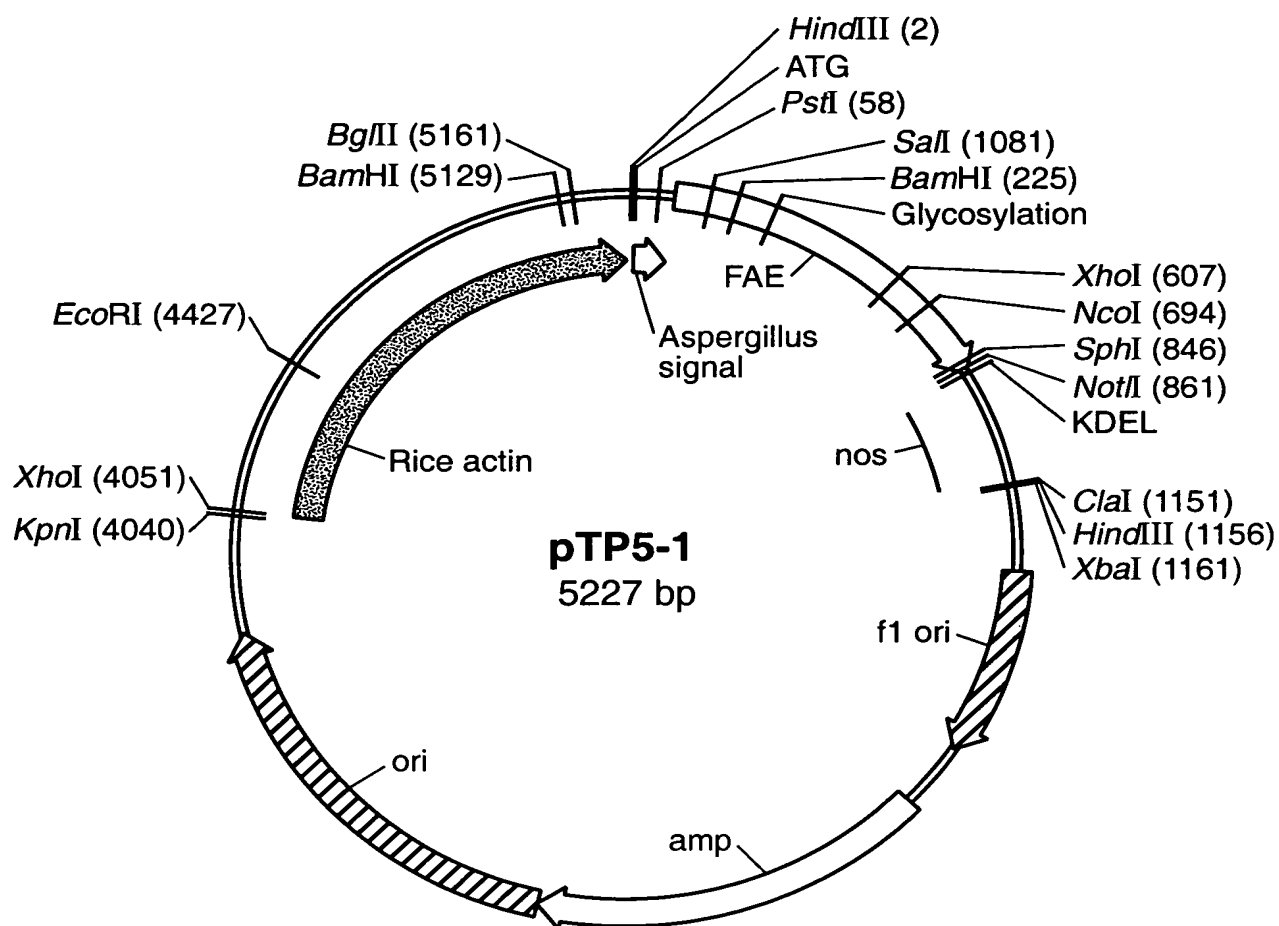


FIG._37A

61 / 154

```
HindIII
~~~~~
      M K Q F S A K H V L A V V V T A G H A L A
1  AAGCTTAACA TGAAGCAGTT CTCCGCCAAA CAGTCTCTCG CAGTTGTGGT GACTGCAGGG CACGCCTTAG
. A S T Q G I S E D L Y S R L V E M A T I S Q A .
71 CAGCCTCTAC GCAAGGCATC TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCACTA TCTCCCAAGC

      Sali
~~~~~
      . A Y A D L C N I P S T I I K G E K I Y N S Q T
141 TGCCTACGCC GACCTGTGCA ACATTCCGTC GACTATTATC AAGGGAGAGA AAATTACAA TTCTCAAACT

      BamHI
~~~~~
      D I N G W I L R D D S S K E I I T V F R G T G S
211 GACATTAACG GATGGATCCT CCGCGACGAC AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCACCTGGTA
. D T N L Q L D T N Y T L T P F D T L P Q C N G .
281 GTGATACGAA TCTACAACTC GATACTAACT ACACCTCTAC GCCTTTCGAC ACCTACCAC AATGCAACGG
. C E V H G G Y I G W V S V Q D Q V E S L V K
351 TTGTGAAGTA CACGGTGGAT ATTATATTGG ATGGGTCTCC GTCCAGGACC AAGTCGAGTC GCTTGTCAAA
Q Q V S Q Y P D Y A L T V T G H X L G A S L A A
421 CAGCAGGTTA GCCAGTATCC GGACTACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG
. L T A A Q L S A T Y D N I R L Y T F G E P R S .
491 CACTCACTGC CGCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC ACCTTCGGCG AACCGCGCAG

      XhoI
~~~~~
      . G N Q A F A S Y M N D A F Q A S S P D T T Q Y
561 CGGCAATCAG GCCTTCGCGT CGTACATGAA CGATGCCCTC CAAGCCTCGA GCCAGATAC GACGCAGTAT

      NcoI
~~~~~
      F R V T H A N D G I P N L P P V E Q G Y A H G G
631 TTCCGGGTCA CTCATGCCAA CGACGGCATC CCAAACCTGC CCCCAGTGA GCAGGGGTAC GCCATGGCG
. V E Y W S V D P Y S A Q N T F V C T G D E V Q .
701 GTGTAGAGTA CTGGAGCGTT GATCCTTACA GCGCCCAGAA CACATTGTCT TGCACCTGGG ATGAAGTGCA
. C C E A Q G G Q G V N N A H T T Y F G M T S G
771 GTGCTGTGAG GCCCAGGGCG GACAGGGTGT GAATAATGCG CACACGACTT ATTTGGGAT GACGAGCGGC
```

FIG._37B

62 / 154

```

SphI      NotI
~~~~~
A  C  T  W  P  V  A  A  A  E  P  L  K  D  E  L  *
841  GCATGCACCT  GGCCGGTTCG  GGCCGCGGAA  CCACTGAAGG  ATGAGCTGTA  AAGAAGCAGA  TCGTTCAAAC
911  ATTTGGCAAT  AAAGTTTCTT  AAGATTGAAT  CCTGTTGCCG  GTCTTGCGAT  GATTATCATA  TAATTTCTGT
981  TGAATTACGT  TAAGCATGTA  ATAATTAAAC  TGTAAATGAT  GACGTTATTT  ATGAGATGGG  TTTTATATGAT
1051 TAGAGTCCCG  CAATTATACA  TTTAATACGC  GATAGAAAC  AAAATATAGC  GCGCAAACTA  GGATAAATTA

HindIII
~~~~~
ClaI      XbaI
~~~~~
1121 TCGCGCGCGG  TGTCATCTAT  GTTACTAGAT  CGATAAGCTT  CTAGAGCGGC  CGGTGGAGCT  CCAATTTCGCC
1191 CTATAGTGAG  TCGTATTACG  CGCGCTCACT  GGCCGTCGTT  TTACAACGTC  GTGACTGGGA  AAACCCCTGGC
1261 GTTACCCCAAC  TTAATCGCCT  TGCAGCACAT  CCCCCTTTCG  CCAGCTGGCG  TAATAGCGAA  GAGGCCCGCA
1331 CCGATCGCCC  TTCCCAACAG  TTGCGCAGCC  TGAATGGCGA  ATGGGACGCG  CCTGTAGCG  GCGCATTAAG
1401 CGCGGCGGGT  GTGGTGGTTA  CGCGCAGCCT  GACCGCTACA  CTTGCCAGCG  CCTAGCGCC  CGCTCCTTTC
1471 GCTTTCCTCC  CTTCCTTTCT  CGCCACGTTT  GCGGCGCTTC  CCCGTCAGC  TCTAAATCGG  GGGCTCCCTT
1541 TAGGGTCCCG  ATTTAGTGCT  TTACGGCAC  TCGACCCCAA  AAACTTGAT  TAGGTGATG  GTTCACGTAG
1611 TGGGCCATCG  CCTGATAGA  CGGTTTTTCG  CCTTTTGACG  TTGGAGTCCA  CGTCTTTAA  TAGTGGACTC
1681 TTGTTCCCAA  CTGGAACAAC  ACTCAACCCT  ATCTCGGTCT  ATTCTTTTGA  TTTATAAGGG  ATTTTGCCGA
1751 TTTTCGGCCTA  TTGGTTAAAA  AATGAGCTGA  TTTAACAAAA  ATTTAACGCG  AATTTTAAAC  AAATATTAAC
1821 GCTTACAAAT  TAGGTGGCAC  TTTTCGGGGA  AATGTGCGCG  GAACCCCTAT  TTGTTTATTT  TTCTAAATAC
1891 ATTCAAATAT  GTATCCGCTC  ATGAGACAT  AACCCTGATA  AATGCTTCAA  TAATATTGAA  AAAGGAAGAG
1961 TATGAGTATT  CAACATTTCC  GTGTGCCCCT  TATTCCTTTT  TTTGCGGCAT  TTTGCTTCC  TGTTTTGTCT
2031 CACCCAGAAA  CGCTGGTGAA  AGTAAAGAT  GCTGAAGATC  AGTTGGGTGC  ACGAGTGGGT  TACATCGAAC
2101 TGGATCTCAA  CAGCGGTAAG  ATCCTTGAGA  GTTTTCGCCC  CGAAGAACGT  TTTCCAATGA  TGAGCACTTT
2171 TAAAGTTCTG  CTATGTGGCG  CGGTATTATC  CCGTATTGAC  GCCGGGCAAG  AGCAACTCGG  TCGCCGCATA
2241 CACTATTCTC  AGAATGACTT  GGTTGAGTAC  TCACCAGTCA  CAGAAAAGCA  TCTTACGGAT  GGCATGACAG
2311 TAAGAGAAAT  ATGCAGTGCT  GCCATAACCA  TGAGTGATAA  CACTGCGGCC  AACTTACTTC  TGACAACGAT
2381 CGGAGGACCG  AAGGAGCTAA  CCGCTTTTTC  GCACAACATG  GGGGATCATG  TAACTCGCCT  TGATCGTTGG
2451 GAACCGGAGC  TGAATGAAGC  CATAACCAAC  GACGAGCGTG  ACACCACGAT  GCCTGTAGCA  ATGGCAACAA
2521 CGTTGCGCAA  ACTATTAACT  GGCGAACTAC  TTACTCTAGC  TTCCCGGCCAA  CAATTAATAG  ACTGGATGGA
2591 GCGGATATAA  GTTGCAGGAC  CACTTCTGCG  CTCGGCCCTT  CCGGCTGGCT  GGTATTATGC  TGATAAATCT
2661 GGAGCCGGTG  AGCGTGGGTC  TCGCGGTATC  ATTGCAGCAC  TGGGGCCAGA  TGGTAAAGCC  TCCCGTATCC
```

FIG._37C

63 / 154

```

2731 TAGTTATCTA CACGACGGGG AGTCAGGCAA CTATGGATGA ACGAAATAGA CAGATCGCTG AGATAGGTGC
2801 CTCACATGATT AAGCATTTGGT AACTGTGCAGA CCAAGTTTAC TCATATATAC TTTAGATTGA TTTAAAACTT
2871 CATTTTAAAT TTAAAAGGAT TTAAGGATGAAG ATCCTTTTGG ATAAATCTCAT GACCAAAATC CTTAAACGTG
2941 AGTTTTCGTT CCACTGAGCG TGCTGCTTGC AACAACAAA AAACCCGCTA CCAGCGGTGG TTTGTTTGCC GGTCTAGTGT
3011 GCGGTAAATC TGCTGCTTGC AACAACAAA AAACCCGCTA CCAGCGGTGG TTTGTTTGCC GGTCTAGTGT
3081 CTACCAATCT TTTTTCGGAA GGTAACTGGC TTCAAGAACT TTCAAGAACT TTCAAGAACT TTCAAGAACT
3151 AGCCGTAGTT AGCCACCCAC GTGCAAGTGC GCGATAAGTC GTGTCTTACC GGTGTTGACT CAAGACGATA GTTACCCGGT
3221 ACCAGTGGCT GCTGCCAGTG GCGATAAGTC GCGATAAGTC GTGTCTTACC GGTGTTGACT CAAGACGATA GTTACCCGGT
3291 AAGGCGAGC GGTCCGGCTG AACGGGGGTG AAGCTATGAG GAGCTATGAG AAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA
3361 AACTGAGATA CCTACAGCGT GGCAGGGTCG GAACAGGAGA GAGCTATGAG AAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA
3431 TCCGGTAAAG GGCAGGGTCG GAACAGGAGA GAGCTATGAG AAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA
3501 TATAGTCCCTG TCGGGTTTCG AACGCGGCTG CTGATTTCTGT GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATACCCGT
3571 GCCTATGGAA AACGCGGCTG CTGATTTCTGT GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATACCCGT
3641 GTTCTTTTCTT GCGTTATCCC GAACGCGGCTG CTGATTTCTGT GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATACCCGT
3711 CGCCGACGCC GAACGCGGCTG CTGATTTCTGT GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATACCCGT
3781 CGCCTCTCCC CGCGCGTTGG CCGATTTCAT AATGCACTAT GCTCACTCAT TAGGCACCCC AGGCTTTACA CTTTATGCTT
3851 CAGTGAGCGC AACGCAATTA ATGTGAGTTA AATGTGAGC GGATAACAAT TTACACACAGG AAACAGCTAT GACCATGATT
3921 CCGGCTCGTA TGTGTGTGG AATGTGAGC GGATAACAAT TTACACACAGG AAACAGCTAT GACCATGATT

3991 ACGCCAAAGC CGCAATTAAC CCTCACTAAA GGGAAACAAA GCTGGGTACC GGGCCCCCCC TCGAGGTGCAT
4061 TCATATGCTT GAGAAAGAG GAGAAAGAGT TCGGGATAGT CCATAATATA ACAAAGGTAA GATTACCTGG TCAAAAAGTGA
4131 AAACATCAGT TAAAAGGTGG TATAAGTAAA ATATCGGTAA TAAAAGGTGG CCAAAAGTGA AATTACTCT
4201 TTTCTACTAT TATAAAAATT GAGGATGTTT TGTCGGTACT TTGATACGTC ATTTTTGTAT GAATTGGTGT
4271 TTAAGTTTAT TCGCGATTGG GAAATGCATA TCTGTATTG AGTCGGTTTT ATAGTTCGTT TAAGTTCGTT GCTTTTGATA
4341 ATACAGAGGG ATTTGTATAA GAAATATCTT TAAAAACCC ATATGCTAAT TTGACATAAT TTTTGAGAAA

4411 AATATATATT CAGGCGAATT CCACAATGAA CAATAATAAG ATTAATAATAG CTTGCCCCCG TTGACAGCGAT
4481 GGGTATTTT TCTAGTAAA TAAAAGATAA ACTTAGACTC AAAACATTTA CAAAACAAAC CCTTAAAGTC
4551 CTAAGGCCCA AAGTGCTATG CACGATCCAT AGCAAGCCCA GCCCAACCCA ACCCAACCCA ACCCAACCCA
4621 GTGCAGCCAA CTGGCAATA GTCTCCACC CCGGCACTAT CACCGTGAGT TGTCCGCACC ACCGCACGTC
4691 TCGCAGCCAA AAAAAAAA AGAAAGAAA AAAAAACAGC AGGTGGGTCC AGGTGGGTCC GGTGCTGGG

```

FIG. 37D

4761 GGCCGGGAAA GCGAGGAGGA TCGCGAGCAG CGACGAGGCC CGGCCCTCCC TCCGCTTCCA AAGAAACGCC
4831 CCCCATCGCC ACTATATACA TACCCCCCCC TCCTCTCCCA TCCCCCAAC CTTACCAACA CCACCAACAC
4901 CACCTCCCTCC CCCCTCGCTG CCGGACGACG AGCTCCCTCC CCTCCCCCT CCGCCGCCGC CGGTAACAC
4971 CCGCCCCCTC TCCTCTTCT TCTCTCGTTT TTTTTCGT CTCGGTCTCG ATCTTTGGCC TTGGTAGTTT
5041 GGGTGGGCGA GAGCGGCTTC GTCGCCCAGA TCGGTGCGCG GGAGGGGCGG GATCTCGCGG CTGGCGTCTC
BamHI
~~~~~  
5111 CGGGCGTGAG TCGGCCCGGA TCCTCGCGGG GAATGGGGCT CTCGGATGTA GATCTCTTTT CTTTCTTCTT  
5181 TTTGTGGTAG AATTGAATC CCTCAGCAAT GTTCATCGGT AGTTTTTCTT TTCATGATTT GTGACAAATG  
5251 CAGCCTCGTG CGGAGCTTTT TTGTAGC  
BglII  
~~~~~

FIG._37E

65 / 154

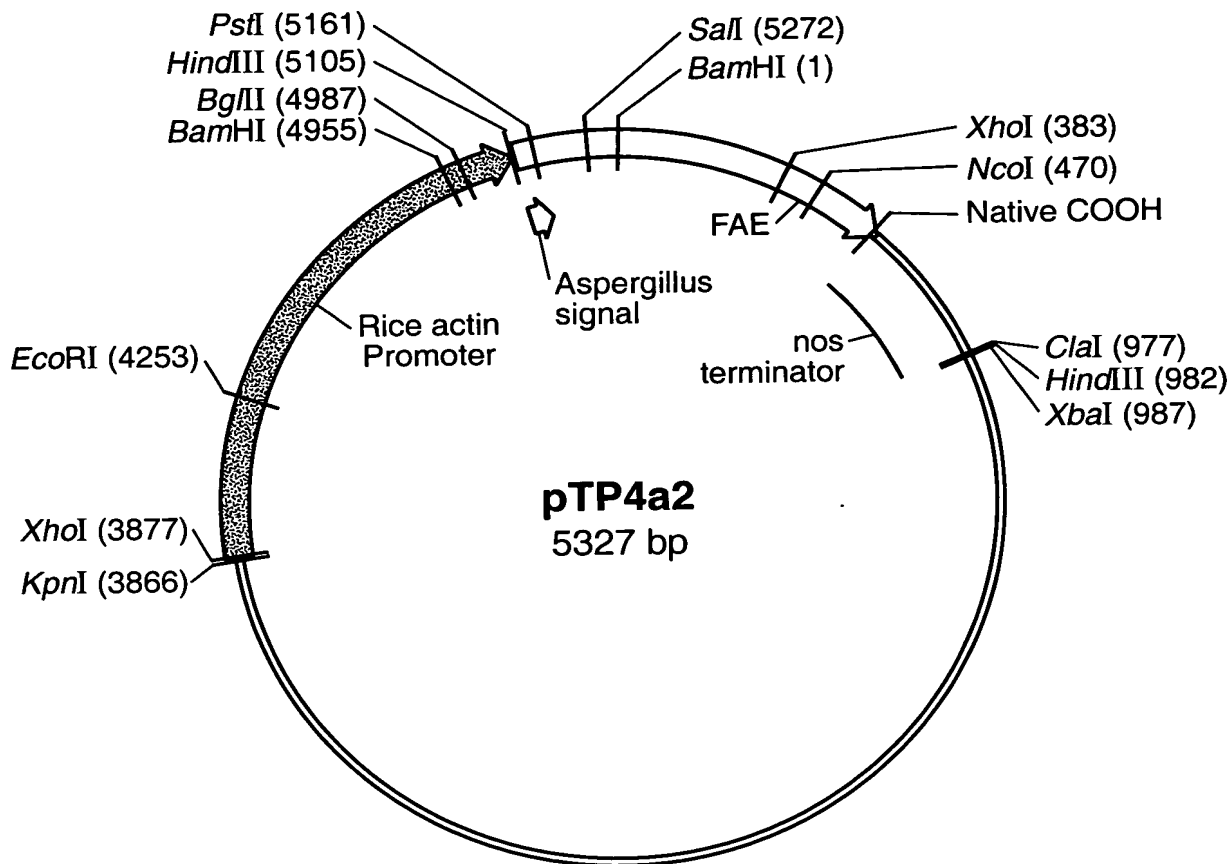


FIG._38A

66 / 154

```
BamHI
~~~~~
. I L R D D S S K E I I T V F R G T G S D T N L
1  GATCCTCCGC GACGACAGCA GCAAAGAAAT AATCACCCTC TTCCGTGGCA CTGGTAGTGA TACGAATCTA
  Q L D T N Y T L T P F D T L P Q C N G C E V H G
71  CAACTCGATA CTAACACTAC CCTCAGCCTT TTCGACACCC TACCACAATG CAACGGTTGT GAAGTACACG
  . G Y Y I G W V S V Q D Q V E S L V K Q Q V S Q .
141 GTGGATATTA TATTGGATGG GTCTCCGTCC AGGACCAAGT CGAGTCGCTT GTCAAAACAGC AGGTTAGCCA
  . Y P D Y A L T V T G H X L G A S L A A L T A A
211 GTATCCGGAC TACGGCTGA CCGTGACCGG CCACKCCCTC GCGCCTCCC TGGCGGCACT CACTGCCGCC
  Q L S A T Y D N I R L Y T F G E P R S G N Q A F
281 CAGCTGTCTG CGACATACGA CAACATCCGC CTGTACACCT TCGGCGAACC GCGCAGCGGC AATCAGGCCT
      xhoI
      ~~~~~
. A S Y M N D A F Q A S S P D T T Q Y F R V T H .
351 TCGCGTCGTA CATGAACGAT GCCTTCCAAG CCTCGAGCCC AGATACGACG CAGTATTTC GGTCACTCA
      NcoI
      ~~~~~
. A N D G I P N L P P V E Q G Y A H G G V E Y W
421 TGCCAACGAC GGCATCCCA ACCTGCCCCC GGTGAGCAG GGTACGCC ATGGCGGTGT AGAGTACTGG
  S V D P Y S A Q N T F V C T G D E V Q C C E A Q
491 AGCGTTGATC CTTACAGCGC CCAGAACACA TTGTCTGCA CTGGGGATGA AGTGCAGTGC TGTGAGGCC
  . G G Q G V N N A H T T Y F G M T S G A C T W * .
561 AGGCGGACA GGGTGTGAAT AATCGCACA CGACTTATTT TGGGATGACG AGCGAGCCT GTACATGGTG
  . *
631 ATCAGTCATT TCAGCCTCCC CGAGTGATCC AGGAAAGATG GATGTCTGG AGAGGGGCC GCGTAACCAC
701 TGAAGGATGA GCTGTAAAGA AGCAGATCGT TCAAACATTT GGCAATAAAG TTTCTTAAGA TTGAATCCCTG
771 TTGCCGGTCT TGCATGATT ATCATATAAT TTCTGTTGAA TTACGTTAAG CATGTAATAA TTAACATGTA
841 ATGCATGACG TTATTATATGA GATGGGTTTT TATGATTAGA GTCCCGCAAT TATACATTTA ATACGCGATA
      ClaI
911 GAAAACAAA TATAGCGCGC AAACATAGGAT AAATTATCGC GCGCGGTGTC ATCTATGTTA CTAGATCGAT
      XbaI
      ~~~~~
HindIII
~~~~~
```

FIG.-38B

981 AAGCTTCTAG AGCGGCCGGT GGAGCTCCAA TTCGCCCTAT AGTGAGTCGT ATTACGGCGG CTCACCTGGCC
1051 GTCGTTTTAC AACGTCGTGA CTGGGAAAC CCTGGCGTTA CCCAACTTAA TCGCCTTGCA GCACATCCCC
1121 CTTTCGCCAG CTGGCGTAAT AGCGAAGAG AGCGCACCGA CCGCCCTTCC CACAGTTGC CACAGCTGAA
1191 TGGCGAATGG CACGCGCCCT GACGCGCCCT GTAGCGGCG ATTAAGCGG CCGGGTGTGG TGGTTACGCG CAGCGTGACC
1261 GCTACACTTG CCAGCGCCCT CACGCGCCCT AGCGCCGCT CTTTCGCTT TCTTCCCTTC CTTTCTCGCC ACGTTCGCCG
1331 GCTTCCCCG TCAAGCTCTA AATCGGGGGC TCCCTTTAGG GTTCCGATTT AGTGCTTTAC GGCACCTCGA
1401 CCCCAGGTTG CTTGATTAGG AGTCCACGTT GTTAAATAGT GGACTCTTGT AGCTAGTGG TCCAAACTGG AACCTATCT
1471 CCGTCTATT CTTTGATTTA TTTTGATTTA TAAGGATTT TGCCGATTTC GGCCTATTGG TTAATAATG AGCTGATTTA
1541 ACAAAATTT AACCGGAATT TTAACAAAAT TTAACGCTT AATATGTAT CCGCTCATGA GACAAATAAC
1611 TCGCGGAAC CCTATTTGT TTAATTTCT AATACATTC GAAGAGTATG TTTGCTCACC TCCTCAACGCT GGTAAAGTA AAAGATGCTG
1681 CTGATAAATG CTTCAATAAT ATTGAAAAAG GAAGAGTATG AATATGTAT AGTATTCATC AATTCCTGT CCGCTTATT
1751 CCCATTTTG CCGCATTTG GGTGCACGA GAACGTTTC CAATGATGAG CACTTTTAA CTTCTGCTAT GTGGCGCGGT ATTATCCCGT
1821 AAGATCAGTT GGTGCACGA GAACGTTTC CAATGATGAG CACTTTTAA CTTCTGCTAT GTGGCGCGGT ATTATCCCGT
1891 TCGCCCCGAA AAGATCAGTT GGTGCACGA GAACGTTTC CAATGATGAG CACTTTTAA CTTCTGCTAT GTGGCGCGGT ATTATCCCGT
1961 ATTGACGCGG AAGATCAGTT GGTGCACGA GAACGTTTC CAATGATGAG CACTTTTAA CTTCTGCTAT GTGGCGCGGT ATTATCCCGT
2031 CAGTCACAGA AAGATCAGTT GGTGCACGA GAACGTTTC CAATGATGAG CACTTTTAA CTTCTGCTAT GTGGCGCGGT ATTATCCCGT
2101 TGATAACACT GCGGCCAAT ATCATGTAAC TCGCCTTGT TACTCTGAC ACGGATGGA AACGATCGGA GGTAAAGTA AAAGATGCTG
2171 AACATGGGG ATCATGTAAC TCGCCTTGT TACTCTGAC ACGGATGGA AACGATCGGA GGTAAAGTA AAAGATGCTG
2241 AGCGTGACAC CACGATGCC TAAATAGACT TAATAGACTG TAGCAATGG CAACAACGTT GATGGAGCG GATAAAGTTG CAGGACCACT TCTGCGCTCG
2311 TCTAGCTTCC CCGCAACAAT TAATAGACTG TAGCAATGG CAACAACGTT GATGGAGCG GATAAAGTTG CAGGACCACT TCTGCGCTCG
2381 CCCCTTCCG CTGGCTGGTT TATTGCTGAT AAGCCCTCCC GTATCGTAGT TATCTACACG ACGGGAGTC AGGCAACTAT
2451 CAGCACTGG GCCAGATGGT AAGCCCTCCC GTATCGTAGT TATCTACACG ACGGGAGTC AGGCAACTAT
2521 GGATGAACGA AATAGACAGA TCGCTGAGAT AGGTGCCCTCA CTGATTAAGC ATTGATAACT GTCAGACCAA
2591 GTTACTCAT ATATACTTTA GATTGATTTA AACTTTTCA TTTAATTTAA AAGGATCTAG GTGAAGATCC
2661 TTTTGTGATA TCTCATGACC AAAATCCCCTT AACGTGAGTT TTCGTTCCAC GTTGCATAAC AAAAACAACA
2731 AAAGATCAAA GGATCTTCTT GAGATCCTTT TTTTCTGCGC GTAATCTGCT CAACCTCTTT TCCGAAGTA ACTGGCTTCA
2801 CCGCTACCAG CCGTGGTTTG TTTGCCGGAT CAAGAGCTAC CAACCTCTTT TCCGAAGTA ACTGGCTTCA
2871 GCAGAGCGCA GATACCAAT ACTGTCTTC TAGTGTAGCC GTAGTTAGGC CACCACTTCA AGAATCTGT
2941 AGCACCGCT ACATACCTCG CTCGTCTAAT CCGGATAAG CGCAGCGGT GAGATACTTA CAGCGTGAGC TATGAGAAAAG
3011 CTTACCGGGT TGGACTCAAG ACGATAGTTA CCGGATAAG CGCAGCGGT GAGATACTTA CAGCGTGAGC TATGAGAAAAG
3081 GCACACAGCC CAGCTTGGAG CGAAGGACCT ACACCGAAT GAGATACTTA CAGCGTGAGC TATGAGAAAAG
3151 CGCCACGCTT CCGGAAGGGA GAAAGGCGGA CAGGTATCCG GTAAAGCGGA GGTTCGGAAC AGGAGAGCGC
3221 ACGAGGGAGC TTCCAGGGG AAACGCTGG TATCTTTATA GTCTGTCTGG GTTTCGCCAC CTCTGACTTG
3291 AGCGTCGATT TTTGTGATGC TCGTCAGGG GCGGAGCCT ATGGAATAAC GCCAGCAACG CGGCCCTTTT

FIG.-38C

FIG. 38D

5111 ACATGAAGCA GTTCTCCGCC AAACACGTCC TCGCAGTTGT GGTGACTGCA GGGCAGCCCT TAGCAGCCTC
5181 TACGCAAGGC ATCTCCGAAG ACCTCTACAG CCGTTAGTC GAAATGGCCA CTATCTCCCA AGCTGCCCTAC
5251 GCCGACCTGT GCAACATTCC GTCGACTATT ATCAAGGGAG AGAAAATTTA CAATTCTCAA ACTGACATTA
5321 ACGGATG

PstI
~~~~~  
M K Q F S A K H V L A V V T A G H A L A A S  
T Q G I S E D L Y S R L V E M A T I S Q A A Y  
TACGCAAGGC ATCTCCGAAG ACCTCTACAG CCGTTAGTC GAAATGGCCA CTATCTCCCA AGCTGCCCTAC  
SalI  
~~~~~  
A D L C N I P S T I I K G E K I Y N S Q T D I N
GCCGACCTGT GCAACATTCC GTCGACTATT ATCAAGGGAG AGAAAATTTA CAATTCTCAA ACTGACATTA
B
~
G W
ACGGATG

FIG._38E

70 / 154

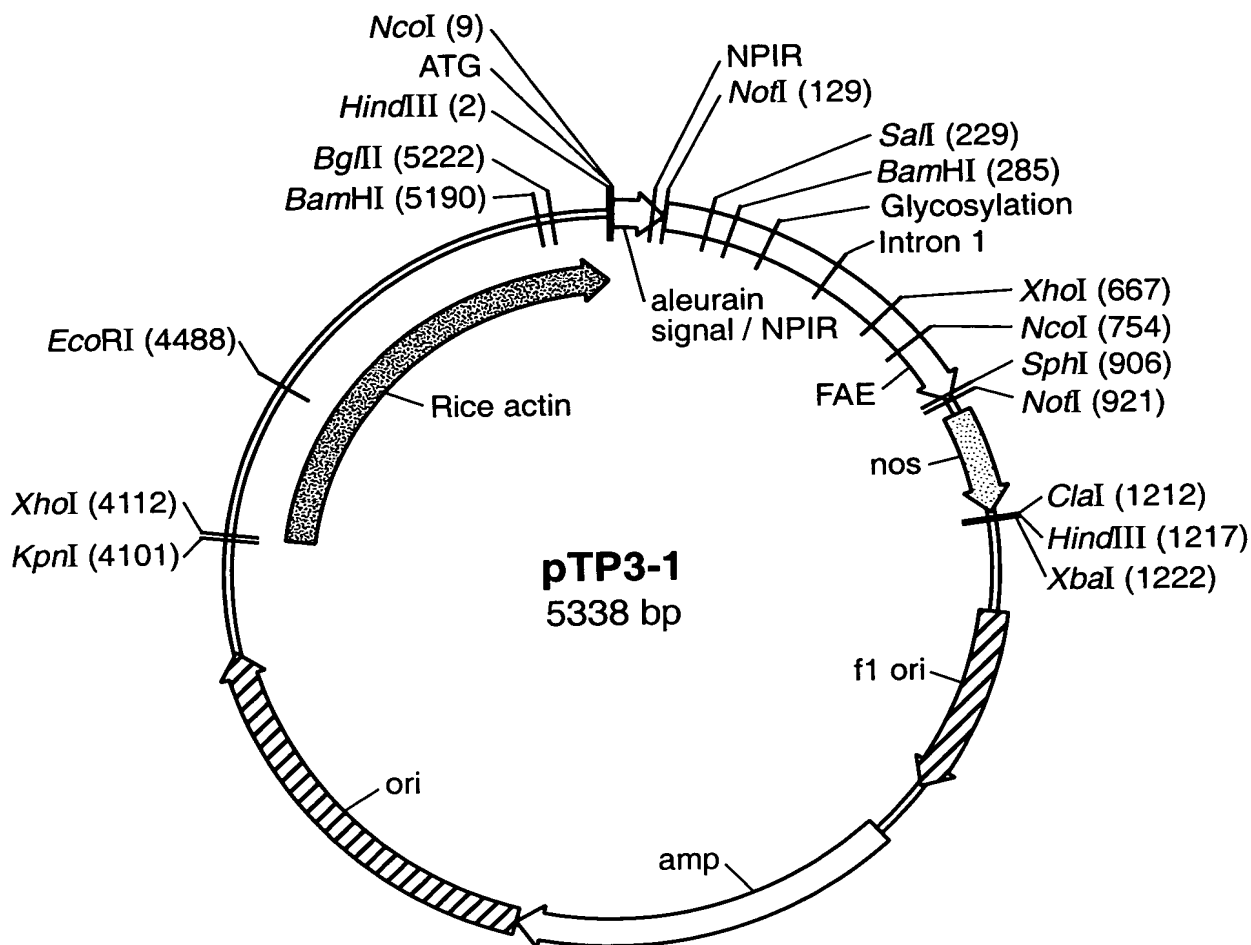


FIG._39A

71 / 154

```

      NcoI
      ~~~~~
HindIII
~~~~~
      M A H A R V L L L A L A V L A T A A V A V
1 AAGCTTACCA TGGCCACGC CCGGTCCTC CTCCTGGCG TCGCCGTGCT GGCACGGCC GCCGTCGCCG

      NPIR
      ~~~~~
      . A S S S F A D S N P I R P V T D R A A A S T .
71 TCGCCTCCTC CTCCTCCTTC GCCGACTCCA ACCGATCCG GCCCGTCACC GACCGCGCGG CCGCCTCCAC
  . Q G I S E D L Y S R L V E M A T I S Q A A Y A
141 GCAGGGCATC TCCGAAGACC TCTACAGCG TTAGTCGAA ATGGCCACTA TCTCCCAAGC TGCCTACGCC

      NotI
      ~~~~~
      . A S S S F A D S N P I R P V T D R A A A S T .
71 TCGCCTCCTC CTCCTCCTTC GCCGACTCCA ACCGATCCG GCCCGTCACC GACCGCGCGG CCGCCTCCAC
  . Q G I S E D L Y S R L V E M A T I S Q A A Y A
141 GCAGGGCATC TCCGAAGACC TCTACAGCG TTAGTCGAA ATGGCCACTA TCTCCCAAGC TGCCTACGCC

      Sali
      ~~~~~
      D L C N I P S T I I K G E K I Y N S Q T D I N G
211 GACCTGTGCA ACATTCCGTC GACTATTATC AAGGGAGAGA AAATTACAA TTCTCAAAC T GACATTAACG

      BamHI
      ~~~~~
      . W I L R D D S S K E I I T V F R G T G S D T N .
281 GATGGATCCT CCGCGACGAC AGCAGCAAAG AAATAATCAC CGTCTCCGT GGCAC TGGTA GTGATACGAA

      Glycosylation
      ~~~~~
      . L Q L D T N Y T L T P F D T L P Q C N G C E V
351 TCTACAAC TC GATACTAACT ACACCCCTCAC GCCTTTCGAC ACCCTACCAC AATGCAACGG TTGTGAAGTA
      H G G Y Y I G W V S V Q D Q V E S L V K Q Q V S
421 CACGGTGGAT ATTATATTGG ATGGGTCTCC GTCCAGGACC AAGTCGAGTC GCTTGTCAAA CAGCAGGTTA
  . Q Y P D Y A L T V T G H X L G A S L A A L T A .
491 GCCAGTATCC GGACTACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG CACTCACTGC
  . A Q L S A T Y D N I R L Y T F G E P R S G N Q
561 CGCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC ACCTTCGGCG AACCGCGCAG CGGCAATCAG
  
```

FIG._39B

72 / 154

```

XhoI
~~~~~
631  A F A S Y M N D A F Q A S S P D T T Q Y F R V T
      GCCTTCGCGT CGTACATGAA CGATGCCTTC CAAGCCTCGA GCCCAGATAC GACGCAGTAT TTCCGGGTCA

NcoI
~~~~~
701  . H A N D G I P N L P P V E Q G Y A H G G V E Y .
      CTCATGCCAA CGACGGCATC CCAAACCTGC CCCCGGTGGA GCAGGGGTAC GCCCATGGCG GTGTAGAGTA
      . W S V D P Y S A Q N T F V C T G D E V Q C C E
      771  CTGGAGCGTT GATCCTTACA GCGCCAGAA CACATTGTC TGCACCTGGG ATGAAGTGCA GTGCTGTGAG

SphI
~~~~~
841  A Q G G Q G V N N A H T T Y F G M T S G A C T W
      GCCCAGGCG GACAGGGTGT GAATAATGCG CACACGACTT ATTTGGGAT GACGAGCGGC GCATGCACCT

NotI
~~~~~
KDEL
~~~~~
911  . P V A A A E T T E G *
      GGCCGGTCGC GGCCGCGGAA ACCACTGAAG GATGAGCTGT AAAGAAGCAG ATCGTTCAA CATTGGCAA
981  TAAAGTTTCT TAAGATTGAA TCCTGTTGCC GGTCTTGCGA TGATTATCAT ATAAATTTCTG TTGAATTACG
1051 TTAAGCATGT AATAATTAA ATGTAATGCA TGACGTTATT TATGAGATGG GTTTTATGA TTAGAGTCCC
1121 GCAATTATAC ATTTAATACG CGATAGAAA CAAATATAG CGCGCAAAC AGGATAAATT ATCGCGCGCG

HindIII
~~~~~
ClaI
~~~~~
XbaI
~~~~~
1191 GTGTCATCTA TGTTACTAGA TCGATAAGCT TCTAGAGCGG CCGGTGGAGC TCCAATTGCG CCTATAGTGA
1261 GTCGTATTAC GCGCGCTCAC TGGCCGTCGT TTTACAACGT CGTGACTGGG AAAACCCCTGG CGTTACCCAA
1331 CTTAATCGCC TTGCAGCACA TCCCCCTTTC GCCAGCTGGC GTAATAGCGA AGAGGCCCGC ACCGATCGCC
1401 CTTCCCAACA GTTGCGCAGC CTGAATGGCG AATGGGACGC GCCCTGTAGC GCGGCATTA GCGCGCGCGG
```

FIG.-39C

73 / 154

1471 TGTGGTGGTT ACGCGCAGCG TGACCGGTAC ACTTGCCAGC GCCCTAGCGC CCGCTCCTTT CGCTTCTTTC
1541 CCTTCCCTTC TCGCCACGTT TCGCCGGCTTT CCGCGTCAAG CTCTAAATCG GGGGCTCCCT TTAGGGTTCC
1611 GATTAGTGC TTACGGCAC CTCGACCCCA AAAAATCTGA TTAGGGTGAT GGTTCACGTA GTGGCCATC
1681 GCCCTGATAG ACGGTTTTC CACCTTTCGCT TATCTCGGTC TATCTTTTG ATTTATAAGG GATTTGCGG ATTTCCGGCT
1751 ACTGGAACAA CACTCAACCC AATGAGCTG AATTAACAA AATTTAAC GAATTTTAA TTTCTAAATA CGCTTACAAT
1821 ATTGGTTAAA CTTTTCGGG CATGAGACAA TAACTCTGAT AAATGCTTCA ATAATAATTGA AAAAGGAAGA GTATGAGTAT
1891 TTAGGTGGCA CTTTTCGGG CATGAGACAA TAACTCTGAT AAATGCTTCA ATAATAATTGA AAAAGGAAGA GTATGAGTAT
1961 TGATACCGCT CATGAGACAA TAACTCTGAT AAATGCTTCA ATAATAATTGA AAAAGGAAGA GTATGAGTAT
2031 TCAACATTC CCGTTCGCC TTAATTCCTT TGCTGAAGAT AGTTTTCGCC CCCGTAATG ATGAGCACAT TTAAGGTTCT
2101 ACGCTGGTGA AAGTAAAGA GATCCTTGAG AGTTTTCGCC CCCGTAATG ATGAGCACAT TTAAGGTTCT
2171 ACAGCGGTAA GATCCTTGAG AGTTTTCGCC CCCGTAATG ATGAGCACAT TTAAGGTTCT
2241 GCTATGTGGC GCGGTATTAT CCCGTAATG ATGAGCACAT TTAAGGTTCT
2311 CAGAAATGACT TGGTTGAGTA CTCACCAAGT ACAGAAAAGC ACAGTGGTG CAGGAGTGG TTTTGCCTTC CTGGATCTCA
2381 TATGCAGTGC TGGCATAACC ACCGCTTTT TGCACAACAT GGGGATCAT GTCCCTGTAG TGCCCTGTAG GACTGGATG GAGCGGATA
2451 GAAGGAGCTA ACCGCTTTT TGCACAACAT GGGGATCAT GTCCCTGTAG TGCCCTGTAG GACTGGATG GAGCGGATA
2521 CTGAATGAAG CCATACCAAA TGGCGAACTA CTTACTCTAG GCTCGGCCCT TCCGCTGGC ACAATTAATA TGCTGTAGT GAGCGGATA
2591 AACTATTAAAC TGGCGAACTA CTTACTCTAG GCTCGGCCCT TCCGCTGGC ACAATTAATA TGCTGTAGT GAGCGGATA
2661 AGTTGCAGGA CCATCTCTGC CTTACTCTAG GCTCGGCCCT TCCGCTGGC ACAATTAATA TGCTGTAGT GAGCGGATA
2731 GAGCGTGGGT CTCGCGGTAT CATTCAGCA ACTATGGATG AACGAAATAG ACAGATCGCT TTTAGATTG ATTTAAACT TCATTTTAA
2801 ACACGACGGG GAGTCAGCA ACTATGGATG AACGAAATAG ACAGATCGCT TTTAGATTG ATTTAAACT TCATTTTAA
2871 TAAGCATTTG TAAGTGTGAG GATCCTTTT GATAATCTCA TGACCAAAAT TTTCTGAGAT TTCTTTTTC TGCGCGTAA
2941 TTTAAAGGA TCTAGGTGAA GTCAGACCC CAAACAAAA AGGTAACTGG AGGTAACTGG AGGTAACTGG AGGTAACTGG
3011 TCCACTGAGC CTGCTGCTTG TAGGCCACCA CTTCAAGAAC GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT
3081 CTGCTGCTTG TAGGCCACCA CTTCAAGAAC GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT
3151 CTTTTCCTGA AGGTAACTGG GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT
3221 TAGGCCACCA CTTCAAGAAC GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT
3291 TGCTGCCAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT
3361 CGGTCGGCT GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT
3431 ACCTACAGG GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT
3501 CGGCAGGTC GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT
3571 GTCGGTTTC GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT GGCATTAAGT
3641 AAAACGCCAG CAACGCGGCC TTTTACGGT TCCTGGCCCT TTGCTGGCTT GATGCTCGTC AGGGGCGG AGCTATGGA
3711 TGCGTTATCC CCTGATTCTG TGGATAACCG TATTAACGCC TTTGAGTGAG TTTGAGTGAG TTTGAGTGAG TTTGAGTGAG
3781 CGAACGACCG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG AAGAGCGCC TTTGAGTGAG TTTGAGTGAG TTTGAGTGAG
3851 CCGCGCGTTG GCCGATTCTAT TAATGCAGCT GGCACGACAG GTTTCCCGAC GTTTCCCGAC GTTTCCCGAC GTTTCCCGAC

FIG._39D

74 / 154

```

3921 CAACGCAATT AATGTGAGTT AGCTCACTCA TTAGGCACCC CAGGCTTTAC ACTTTATGCT TCCGGCTCGT
3991 ATGTTGTGTG GAATTGTGAG CGGATAACAA TTTACACACAG GAAACAGCTA TGACCATGAT TACGCCAAGC
                                     KpnI
                                     ~~~~~
4061 GCGCAATTAA CCTTCACTAA AGGGAACAAA AGCTGGGTAC CGGGCCCCCC CTCGAGGTCA TTCATATGCT
4131 TGAGAAGAGA GTCGGGATAG TCCAAAATAA AACAAAGTA AGATTACCTG GTCAAAAGTG AAAACATCAG
4201 TTTAAAGGTG GTATAAGTAA AATATCGGTA ATAAAGGTG GCCCAAAGTG AAATTACTC TTTTCTACTA
4271 TTATAAAAAT TGAGGATGTT TTGTCGGTAC TTTGATACGT CATTTTGTG TGAATTGGTT TTTAAGTTTA
4341 TTCGCGATTT GGAATGTCAT ATCTGTATTT GAGTCGGTTT TTAAGTTTCT TGTTTTGTG AATACAGAGG
4411 GATTGTATA AGAAATATCT TTAATAAACCAT CATATGCTAA TTTGACATAA TTTTGTGAGAA AAATATATAT
                                     EcoRI
                                     ~~~~~
4481 TCAGGCGAAT TCCACAATGA ACAATAATAA GATTAAATA GCTTGCCCCC GTTGCAGCGA TGGGTATTTT
4551 TTCTAGTAA ATAAAGATA AACTTAGACT CAAAACATTT ACAAAAACAA CCCCTAAAGT CCTAAAGCCC
4621 AAAGTGCTAT GCACGATCCA TAGCAAGCCC AGCCCAACCC AACCCACCCC AACCCACCCC AGTGCAGCCA
4691 ACTGGCAAAT AGTCTCCACC CCGGGCACTA TCACCGTGAG TTGTCCGCAC CACCGCACGT CTCGCAGCCA
4761 AAAAAAAA AAGAAAGAAA AAAAAAGAAA AGAAACACAG CAGGTGGGTC CGGGTCGTGG GGGCCGGAAA
4831 AGCGAGGAGG ATCGCGAGCA GCGACGAGGC CCGGCCCTCC CTCCGCTTCC AAAGAAACGC CCCCCATCGC
4901 CACTATATAC ATACCCCCCC CTCTCCTCCC ATCCCCCCTCC CCCTACCAAC ACCACCAACA CCACCTCCTC
4971 CCCCCTCGCT GCCGGACGAC GAGCTCCTCC CCCCCTCCCC TCCGCCGCCG CCGGTAACCA CCCCCCCCCC
5041 CTCCTCTTTC TTTCTCCGTT TTTTCTTTCG TCTCGGTCTC GATCTTTGGC CTTGGTAGTT TGGGTGGGCG
5111 AGAGCGGCTT CGTCGCCCAG ATCGGTGCGC GGGAGGGGCG GGATCTCGCG GCTGGCGTCT CCGGGCGGTGA
                                     BamHI
                                     ~~~~~
5181 GTCGGCCCCG ATCCTCGCGG GGAATGGGCG TCTCGGATGT AGATCTTCTT TCTTCTTCTT TTTTGTGGTA
5251 GAAATTGAAT CCTCAGCAT TGTTCATCGG TAGTTTTTCT TTTTCATGATT TGTGACAAAT GCAGCCTCGT
5321 GCGGAGCTTT TTTGTAGC

```

FIG._39E

75 / 154

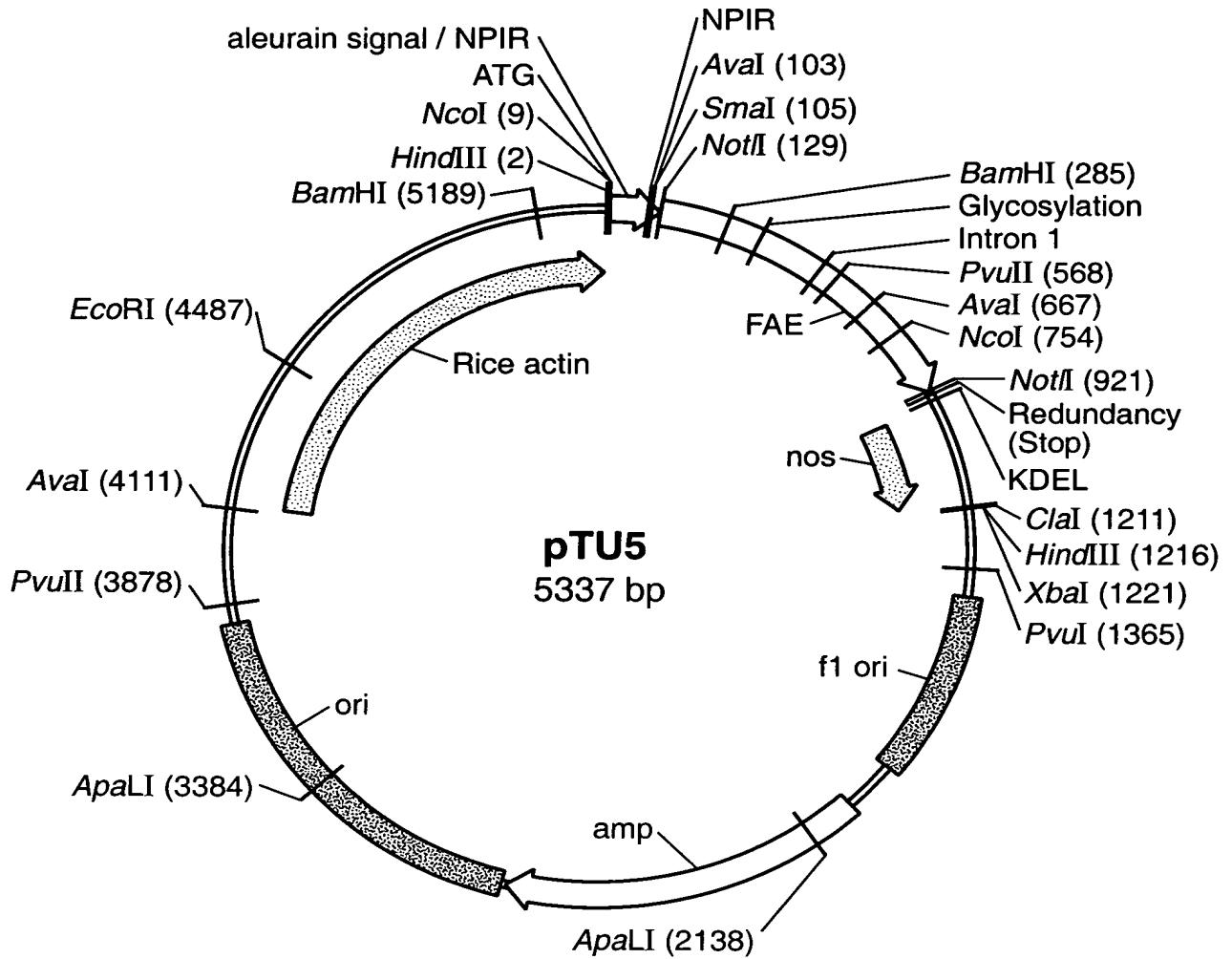


FIG._40A

76 / 154

HindIII NcoI
~~~~~

1 AAGCTTACCA TGGCCCACGC CCGCGTCCTC CTCCTGGCGC TCGCCGTGCT  
TTCGAATGGT ACCGGGTGCG GGCGCAGGAG GAGGACCGCG AGCGGCACGA

51 GGCCACGGCC GCCGTCGCCG TCGCCTCCTC CTCCTCCTTC GCCGACTCCA  
CCGGTGCCGG CGGCAGCGGC AGCGGAGGAG GAGGAGGAAG CGGCTGAGGT

**SmaI**  
~~~~~

AvaI
~~~~~

**NotI**  
~~~~~

101 ACCCGGGCCG GCCCGTCACC GACCGCGCGG CCGCCTCCAC GCAGGGCATC
TGGGCCCGGC CGGGCAGTGG CTGGCGCGCC GGCGGAGGTG CGTCCCGTAG

151 TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCACTA TCTCCCAAGC
AGGCTTCTGG AGATGTCGGC AAATCAGCTT TACCGGTGAT AGAGGGTTCG

201 TGCCTACGCC GACCTGTGCA ACATTCCGTC GACTATTATC AAGGGAGAGA
ACGGATGCGG CTGGACACGT TGTAAGGCAG CTGATAATAG TTCCCTCTCT

BamHI
~~~~~

251 AAATTTACAA TTCTCAAAC TACATTAACG GATGGATCCT CCGCGACGAC  
TTTAAATGTT AAGAGTTTGA CTGTAATTGC CTACCTAGGA GGCCTGCTG

301 AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCACCTGGTA GTGATACGAA  
TCGTCGTTTC TTTATTAGTG GCAGAAGGCA CCGTGACCAT CACTATGCTT

351 TCTACAAC TC GATACTAACT ACACCCTCAC GCCTTTTCGAC ACCCTACCAC  
AGATGTTGAG CTATGATTGA TGTGGGAGTG CGGAAAGCTG TGGGATGGTG

401 AATGCAACGG TTGTGAAGTA CACGGTGGAT ATTATATTGG ATGGGTCTCC  
TTACGTTGCC AACACTTCAT GTGCCACCTA TAATATAACC TACCCAGAGG

451 GTCCAGGACC AAGTCGAGTC GCTTGTCAAA CAGCAGGTTA GCCAGTATCC  
CAGGTCTTGG TTCAGCTCAG CGAACAGTTT GTCGTCCAAT CGGTCATAGG

501 GGA CTACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG  
CCTGATGCGC GACTGGCACT GGCCGGTGMG GGAGCCGCGG AGGGACCGCC

**PvuII**  
~~~~~

551 CACTCACTGC CGCCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC
GTGAGTGACG GCGGGTCGAC AGACGCTGTA TGCTGTTGTA GGC GGACATG

601 ACCTTCGGCG AACCGCGCAG CGGCAATCAG GCCTTCGCGT CGTACATGAA
TGGAAGCCGC TTGGCGCGTC GCCGTTAGTC CGGAAGCGCA GCATGTACTT

AvaI
~~~~~

651 CGATGCCTTC CAAGCCTCGA GCCCAGATAC GACGCAGTAT TTCCGGGTCA  
GCTACGGAAG GTTCGGAGCT CGGGTCTATG CTGCGTCATA AAGGCCAGT

**FIG.\_40B**

77 / 154

701 CTCATGCCAA CGACGGCATC CCAAACCTGC CCCCGGTGGA GCAGGGGTAC  
GAGTACGGTT GCTGCCGTAG GGTTTGGACG GGGGCCACCT CGTCCCCATG

NcoI

~~~~~

751 GCCCATGGCG GTGTAGAGTA CTGGAGCGTT GATCCTTACA GCGCCCAGAA
CGGGTACCGC CACATCTCAT GACCTCGCAA CTAGGAATGT CGCGGGTCTT

801 CACATTTGTC TGCCTGGGG ATGAAGTGCA GTGCTGTGAG GCCCAGGGCG
GTGTAAACAG ACGTGACCCC TACTTCACGT CACGACACTC CGGGTCCCGC

851 GACAGGGTGT GAATAATGCG CACACGACTT ATTTTGGGAT GACGAGCGGC
CTGTCCCACA CTTATTACGC GTGTGCTGAA TAAAACCCTA CTGCTCGCCG

NotI

~~~~~

901 GCATGCACCT GGCCGGTCGC GGCCGCGGAA CCACTGAAGG ATGAGCTGTA  
CGTACGTGGA CCGGCCAGCG CCGGCGCCTT GGTGACTTCC TACTCGACAT

951 AAGAAGCAGA TCGTTCAAAC ATTTGGCAAT AAAGTTTCTT AAGATTGAAT  
TTCTTCGTCT AGCAAGTTTG TAAACCGTTA TTTCAAAGAA TTCTAACTTA

1001 CCTGTTGCCG GTCTTGCGAT GATTATCATA TAATTTCTGT TGAATTACGT  
GGACAACGGC CAGAACGCTA CTAATAGTAT ATTAAAGACA ACTTAATGCA

1051 TAAGCATGTA ATAATTAACA TGTAATGCAT GACGTTATTT ATGAGATGGG  
ATTCGTACAT TATTAATTGT ACATTACGTA CTGCAATAAA TACTCTACCC

1101 TTTTATGAT TAGAGTCCCG CAATTATACA TTTAATACGC GATAGAAAAC  
AAAAATACTA ATCTCAGGGC GTTAATATGT AAATTATGCG CTATCTTTTG

1151 AAAATATAGC GCGCAAACCTA GGATAAATTA TCGCGCGCGG TGTATCTAT  
TTTTATATCG CGCGTTTGAT CCTATTTAAT AGCGCGCGCC ACAGTAGATA

XbaI

~~~~~

ClaI HindIII

~~~~~

1201 GTTACTAGAT CGATAAGCTT CTAGAGCGGC CGGTGGAGCT CCAATTCGCC  
CAATGATCTA GCTATTCGAA GATCTCGCCG GCCACCTCGA GGTAAAGCGG

1251 CTATAGTGAG TCGTATTACG CGCGCTCACT GGCCGTCGTT TTACAACGTC  
GATATCACTC AGCATAATGC GCGCGAGTGA CCGGCAGCAA AATGTTGCAG

1301 GTGACTGGGA AAACCCTGGC GTTACCCAAC TTAATCGCCT TGCAGCACAT  
CACTGACCCT TTTGGGACCG CAATGGGTTG AATTAGCGGA ACGTCGTGTA

PvuII

~~~~~

1351 CCCCTTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC
GGGGGAAAGC GGTCGACCGC ATTATCGCTT CTCCGGGCGT GGCTAGCGGG

1401 TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGGACGCG CCCTGTAGCG
AAGGGTTGTC AACGCGTCGG ACTTACCGCT TACCCTGCGC GGGACATCGC

FIG. 40C

78 / 154

1451	GCGCATTAAG CGCGTAATTC	CGCGGCGGGT GCGCCGCCCA	GTGGTGGTTA CACCACCAAT	CGCGCAGCGT GCGCGTCGCA	GACCGCTACA CTGGCGATGT
1501	CTTGCCAGCG GAACGGTCGC	CCCTAGCGCC GGGATCGCGG	CGCTCCTTTC GCGAGGAAAG	GCTTTCTTCC CGAAAGAAGG	CTTCCTTTCT GAAGGAAAGA
1551	CGCCACGTTT GCGGTGCAAG	GCCGGCTTTC CGGCCGAAAG	CCCGTCAAGC GGGCAGTTCG	TCTAAATCGG AGATTTAGCC	GGGCTCCCTT CCCGAGGGAA
1601	TAGGGTTCCG ATCCCAAGGC	ATTTAGTGCT TAAATCACGA	TTACGGCACC AATGCCGTGG	TCGACCCCAA AGCTGGGGTT	AAAAC TTGAT TTTTGAACTA
1651	TAGGGTGATG ATCCCACTAC	G TTCACGTAG CAAGTGCATC	TGGGCCATCG ACCCGGTAGC	CCCTGATAGA GGGACTATCT	CGGTTTTTTCG GCCAAAAAGC
1701	CCCTTTGACG GGGAAACTGC	TTGGAGTCCA AACCTCAGGT	CGTTCTTTAA GCAAGAAATT	TAGTGGACTC ATCACCTGAG	TTGTTCCAAA AACAAGGTTT
1751	CTGGAACAAC GACCTTGTTG	ACTCAACCCT TGAGTTGGGA	ATCTCGGTCT TAGAGCCAGA	ATTCTTTTGA TAAGAAAAC T	TTTATAAGGG AAATATTCCC
1801	ATTTTGCCGA TAAAACGGCT	TTTCGGCCTA AAAGCCGGAT	TTGGTTAAAA AACCAATTTT	AATGAGCTGA TTACTCGACT	TTTAACAAAA AAATTGTTTT
1851	ATTTAACGCG TAAATTGCGC	AATTTTAAAC TTAAAATTGT	AAATATTAAC TTTATAATTG	GCTTACAATT CGAATGT TAA	TAGGTGGCAC ATCCACCGTG
1901	TTTTCGGGGA AAAAGCCCCT	AATGTGCGCG TTACACGCGC	GAACCCCTAT CTTGGGGATA	TTGTTTATTT AACAAATAAA	TTCTAAATAC AAGATTTATG
1951	ATTCAAATAT TAAGTTTATA	GTATCCGCTC CATAGGCGAG	ATGAGACAAT TACTCTGTTA	AACCCTGATA TTGGGACTAT	AATGCTTCAA TTACGAAGTT
2001	TAATATTGAA ATTATAACTT	AAAGGAAGAG TTTCCTTCTC	TATGAGTATT ATACTCATAA	CAACATTTCC GTTGTAAAGG	GTGTCGCCCT CACAGCGGGA
2051	TATTCCTTTT ATAAGGGAAA	TTTGCGGCAT AAACGCCGTA	TTTGCCTTCC AAACGGAAGG	TGTTTTTGCT ACAAAAACGA	CACCCAGAAA GTGGGTCTTT

ApaLI

~~~~~

|      |                          |                          |                           |                          |                          |
|------|--------------------------|--------------------------|---------------------------|--------------------------|--------------------------|
| 2101 | CGCTGGTGAA<br>GCGACCACTT | AGTAAAAGAT<br>TCATTTTCTA | GCTGAAGATC<br>CGACTTCTAG  | AGTTGGGTGC<br>TCAACCACG  | ACGAGTGGGT<br>TGCTCACCCA |
| 2151 | TACATCGAAC<br>ATGTAGCTTG | TGGATCTCAA<br>ACCTAGAGTT | CAGCGGTAAG<br>GTCGCCATTC  | ATCCTTGAGA<br>TAGGAACTCT | GTTTTCGCCC<br>CAAAAGCGGG |
| 2201 | CGAAGAACGT<br>GCTTCTTGCA | TTTCCAATGA<br>AAAGGTTACT | TGAGCACTTT<br>ACTCGTGAAA  | TAAAGTTCTG<br>ATTTCAAGAC | CTATGTGGCG<br>GATACACCGC |
| 2251 | CGGTATTATC<br>GCCATAATAG | CCGTATTGAC<br>GGCATAACTG | GCCGGGCAAG<br>CGGCCCCGTT  | AGCAACTCGG<br>TCGTTGAGCC | TCGCCGCATA<br>AGCGGCGTAT |
| 2301 | CACTATTCTC<br>GTGATAAGAG | AGAATGACTT<br>TCTTACTGAA | GGTTGAGTAC<br>CCAAC TCATG | TCACCAGTCA<br>AGTGGTCAGT | CAGAAAAGCA<br>GTCTTTTCTG |

FIG.\_40D

79 / 154

|      |                           |                          |                          |                           |                          |
|------|---------------------------|--------------------------|--------------------------|---------------------------|--------------------------|
| 2351 | TCTTACGGAT<br>AGAATGCCTA  | GGCATGACAG<br>CCGTACTGTC | TAAGAGAATT<br>ATTCTCTTAA | ATGCAGTGCT<br>TACGTCACGA  | GCCATAACCA<br>CGGTATTGGT |
| 2401 | TGAGTGATAA<br>ACTCACTATT  | CACTGCGGCC<br>GTGACGCCGG | AACTTACTTC<br>TTGAATGAAG | TGACAACGAT<br>ACTGTTGCTA  | CGGAGGACCG<br>GCCTCCTGGC |
| 2451 | AAGGAGCTAA<br>TTCCTCGATT  | CCGCTTTTTT<br>GGCGAAAAAA | GCACAACATG<br>CGTGTGTGAC | GGGGATCATG<br>CCCCTAGTAC  | TAACTCGCCT<br>ATTGAGCGGA |
| 2501 | TGATCGTTGG<br>ACTAGCAACC  | GAACCGGAGC<br>CTTGGCCTCG | TGAATGAAGC<br>ACTTACTTCG | CATACCAAAC<br>GTATGGTTTG  | GACGAGCGTG<br>CTGCTCGCAC |
| 2551 | ACACCACGAT<br>TGTGGTGCTA  | GCCTGTAGCA<br>CGGACATCGT | ATGGCAACAA<br>TACCGTTGTT | CGTTGCGCAA<br>GCAACGCGTT  | ACTATTAAC<br>TGATAATTGA  |
| 2601 | GGCGAACTAC<br>CCGCTTGATG  | TTACTCTAGC<br>AATGAGATCG | TTCCCGGCAA<br>AAGGGCCGTT | CAATTAATAG<br>GTTAATTATC  | ACTGGATGGA<br>TGACCTACCT |
| 2651 | GGCGGATAAA<br>CCGCCTATTT  | GTTGCAGGAC<br>CAACGTCTTG | CACTTCTGCG<br>GTGAAGACGC | CTCGGCCCTT<br>GAGCCGGGAA  | CCGGCTGGCT<br>GGCCGACCGA |
| 2701 | GGTTTATTGC<br>CCAAATAACG  | TGATAAATCT<br>ACTATTTAGA | GGAGCCGGTG<br>CCTCGGCCAC | AGCGTGGGTC<br>TCGCACCCAG  | TCGCGGTATC<br>AGCGCCATAG |
| 2751 | ATTGCAGCAC<br>TAACGTCGTG  | TGGGGCCAGA<br>ACCCCGGTCT | TGGTAAGCCC<br>ACCATTCCGG | TCCCGTATCG<br>AGGGCATAGC  | TAGTTATCTA<br>ATCAATAGAT |
| 2801 | CACGACGGGG<br>GTGCTGCCCC  | AGTCAGGCAA<br>TCAGTCCGTT | CTATGGATGA<br>GATACCTACT | ACGAAATAGA<br>TGCTTTATCT  | CAGATCGCTG<br>GTCTAGCGAC |
| 2851 | AGATAGGTGC<br>TCTATCCACG  | CTCACTGATT<br>GAGTGACTAA | AAGCATTGGT<br>TTCGTAACCA | AACTGTCAGA<br>TTGACAGTCT  | CCAAGTTTAC<br>GGTTCAAATG |
| 2901 | TCATATATAC<br>AGTATATATG  | TTTAGATTGA<br>AAATCTAACT | TTTAAACTTT<br>AAATTTTGAA | CATTTTTTAAT<br>GTAAAAATTA | TTAAAGGAT<br>AATTTTCCTA  |
| 2951 | CTAGGTGAAG<br>GATCCACTTC  | ATCCTTTTTG<br>TAGGAAAAAC | ATAATCTCAT<br>TATTAGAGTA | GACCAAATC<br>CTGGTTTTAG   | CCTTAACGTG<br>GGAATTGCAC |
| 3001 | AGTTTTTCGTT<br>TCAAAAGCAA | CCACTGAGCG<br>GGTGACTCGC | TCAGACCCCG<br>AGTCTGGGGC | TAGAAAAGAT<br>ATCTTTTCTA  | CAAAGGATCT<br>GTTTCCTAGA |
| 3051 | TCTTGAGATC<br>AGAACTCTAG  | CTTTTTTTCT<br>GAAAAAAGA  | GCGCGTAATC<br>CGCGCATTAG | TGCTGCTTGC<br>ACGACGAACG  | AAACAAAAAA<br>TTTGTTTTTT |
| 3101 | ACCACCGCTA<br>TGGTGGCGAT  | CCAGCGGTGG<br>GGTCGCCACC | TTTGTGTGCC<br>AAACAAACGG | GGATCAAGAG<br>CCTAGTTCTC  | CTACCAACTC<br>GATGGTTGAG |
| 3151 | TTTTTCCGAA<br>AAAAAGGCTT  | GGTAACTGGC<br>CCATTGACCG | TTCAGCAGAG<br>AAGTCGTCTC | CGCAGATACC<br>GCGTCTATGG  | AAATACTGTC<br>TTTATGACAG |
| 3201 | CTTCTAGTGT<br>GAAGATCACA  | AGCCGTAGTT<br>TCGGCATCAA | AGGCCACCAC<br>TCCGGTGGTG | TTCAAGAACT<br>AAGTTCTTGA  | CTGTAGCACC<br>GACATCGTGG |
| 3251 | GCCTACATAC<br>CGGATGTATG  | CTCGCTCTGC<br>GAGCGAGACG | TAATCCTGTT<br>ATTAGGACAA | ACCAGTGGCT<br>TGGTCACCGA  | GCTGCCAGTG<br>CGACGGTCAC |

**FIG. 40E**

80 / 154

3301 GCGATAAGTC GTGTCTTACC GGGTTGGACT CAAGACGATA GTTACCGGAT  
CGCTATTACG CACAGAATGG CCCAACCTGA GTTCTGCTAT CAATGGCCTA

ApaLI

~~~~~

3351 AAGGCGCAGC GGTCGGGCTG AACGGGGGGT TCGTGACAC AGCCAGCTT
TTCCGCGTCG CCAGCCCGAC TTGCCCCCA AGCACGTGTG TCGGGTCGAA

3401 GGAGCGAACG ACCTACACCG AACTGAGATA CCTACAGCGT GAGCTATGAG
CCTCGCTTGC TGGATGTGGC TTGACTCTAT GGATGTGCA CTCGATACTC

3451 AAAGCGCCAC GCTTCCCGAA GGGAGAAAG CGGACAGGTA TCCGGTAAGC
TTTCGCGGTG CGAAGGGCTT CCCTCTTTCC GCCTGTCCAT AGGCCATTGC

3501 GGCAGGGTCG GAACAGGAGA GCGCACGAGG GAGCTTCCAG GGGGAAACGC
CCGTCCAGC CTTGTCCTCT CGCGTGCTCC CTCGAAGGTC CCCCTTTGCG

3551 CTGGTATCTT TATAGTCTTG TCGGGTTTCG CCACCTCTGA CTTGAGCGTC
GACCATAGAA ATATCAGGAC AGCCCAAAGC GGTGGAGACT GAACTCGCAG

3601 GATTTTTGTG ATGCTCGTCA GGGGGGCGGA GCCTATGGAA AAACGCCAGC
CTAAAAACAC TACGAGCAGT CCCCCGCCT CGGATACCTT TTTGCGGTGC

3651 AACGCGGCCT TTTTACGGTT CCTGGCCTTT TGCTGGCCTT TTGCTCACAT
TTGCGCCGGA AAAATGCCAA GGACCGGAAA ACGACCGGAA AACGAGTGTA

3701 GTTCTTTCCT GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCCT
CAAGAAAGGA CGCAATAGGG GACTAAGACA CCTATTGGCA TAATGGCGGA

3751 TTGAGTGAGC TGATACCGCT CGCCGCAGCC GAACGACCGA GCGCAGCGAG
AACTCACTCG ACTATGGCGA GCGGCGTCGG CTTGCTGGCT CGCGTCGCTC

3801 TCAGTGAGCG AGGAAGCGGA AGAGCGCCCA ATACGCAAAC CGCCTCTCCC
AGTCACTCGC TCCTTCGCCT TCTCGCGGGT TATGCGTTTG GCGGAGAGGG

PvuII

~~~~~

3851 CGCGCGTTGG CCGATTCATT AATGCAGCTG GCACGACAGG TTTCCCGACT  
GCGCGCAACC GGCTAAGTAA TTACGTCGAC CGTGCTGTCC AAAGGGCTGA

3901 GGAAAGCGGG CAGTGAGCGC AACGCAATTA ATGTGAGTTA GCTCACTCAT  
CCTTTCGCCC GTCACCTCGC TTGCGTTAAT TACACTCAAT CGAGTGAGTA

3951 TAGGCACCCC AGGCTTTACA CTTTATGCTT CCGGCTCGTA TGTGTGTGG  
ATCCGTGGGG TCCGAAATGT GAAATACGAA GGCCGAGCAT ACAACACACC

4001 AATTGTGAGC GGATAACAAT TTCACACAGG AAACAGCTAT GACCATGATT  
TTAACACTCG CCTATTGTGA AAGTGTGTCC TTTGTCGATA CTGGTACTAA

4051 ACGCCAAGCG CGCAATTAAC CCTCACTAAA GGGAACAAAA GCTGGGTACC  
TGCGGTTCGC GCGTTAATTG GGAGTGATTT CCCTTGTTTT CGACCCATGG

AvaI

~~~~~

4101 GGGCCCCCCC TCGAGGTCAT TCATATGCTT GAGAAGAGAG TCGGGATAGT
CCCGGGGGGG AGCTCCAGTA AGTATACGAA CTCTTCTCTC AGCCCTATCA

FIG. 40F

81 / 154

4151	CCAAAATAAA GGTTTTATTT	ACAAAGGTAA TGTTTCCATT	GATTACCTGG CTAATGGACC	TCAAAAGTGA AGTTTTCACT	AAACATCAGT TTTGTAGTCA
4201	TAAAAGGTGG ATTTTCCACC	TATAAGTAAA ATATTCATTT	ATATCGGTAA TATAGCCATT	TAAAAGGTGG ATTTTCCACC	CCCAAAGTGA GGGTTTCACT
4251	AATTTACTCT TTAAATGAGA	TTTCTACTAT AAAGATGATA	TATAAAAATT ATATTTTTAA	GAGGATGTTT CTCCTACAAA	TGTCGGTACT ACAGCCATGA
4301	TTGATACGTC AACTATGCAG	ATTTTTGTAT TAAAAACATA	GAATTGGTTT CTTAACCAAA	TTAAGTTTAT AATTCAAATA	TCGCGATTTG AGCGCTAAAC
4351	GAAATGCATA CTTTACGTAT	TCTGTATTTG AGACATAAAC	AGTCGGTTTT TCAGCCAAAA	TAAAGTTCGT ATTCAAGCAA	GCTTTTGTAA CGAAAACATT
4401	ATACAGAGGG TATGTCTCCC	ATTTGTATAA TAAACATATT	GAAATATCTT CTTTATAGAA	TAAAAAACCC ATTTTTTGGG	ATATGCTAAT TATACGATTA
EcoRI ~~~~~					
4451	TTGACATAAT AACTGTATTA	TTTTGAGAAA AAAACCTCTT	AATATATATT TTATATATAA	CAGGCGAATT GTCCGCTTAA	CCACAATGAA GGTGTTACTT
4501	CAATAATAAG GTTATTATTC	ATTAAAATAG TAATTTTATC	CTTGCCCCCG GAACGGGGGC	TTGCAGCGAT AACGTCGCTA	GGGTATTTTT CCCATAAAAA
4551	TCTAGTAAAA AGATCATTTT	TAAAAGATAA ATTTTCTATT	ACTTAGACTC TGAATCTGAG	AAAACATTTA TTTTGTAAAT	CAAAAACAAC GTTTTTGTG
4601	CCCTAAAGTC GGGATTTTCAG	CTAAAGCCCA GATTTTCGGT	AAGTGCTATG TTCACGATAC	CACGATCCAT GTGCTAGGTA	AGCAAGCCCA TCGTTCGGGT
4651	GCCCAACCCA CGGGTTGGGT	ACCCAACCCA TGGGTGTTGGT	ACCCACCCCA TGGGTGTTGGT	GTGCAGCCAA CACGTCGGTT	CTGGCAAATA GACCGTTTAT
4701	GTCTCCACCC CAGAGGTGGG	CCGGCACTAT GGCCGTGATA	CACCGTGAGT GTGGCACTCA	TGTCCGCACC ACAGGCGTGG	ACCGCACGTC TGGCGTGCAG
4751	TCGCAGCCAA AGCGTCGGTT	AAAAAAAAAA TTTTTTTTTT	AGAAAGAAAA TCTTCTTTTT	AAAAGAAAAA TTTTCTTTTT	GAAAAACAGC CTTTTTGTGC
4801	AGGTGGGTCC TCCACCCAGG	GGGTCGTGGG CCCAGCACCC	GGCCGGAAAA CCGGCCTTTT	GCGAGGAGGA CGCTCCTCCT	TCGCGAGCAG AGCGCTCGTC
4851	CGACGAGGCC GCTGCTCCGG	CGGCCCTCCC GCCGGGAGGG	TCCGCTTCCA AGGCGAAGGT	AAGAAACGCC TTCCTTTGCGG	CCCCATCGCC GGGGTAGCGG
4901	ACTATATACA TGATATATGT	TACCCCCCCC ATGGGGGGGG	TCTCCTCCCA AGAGGAGGGT	TCCCCCAAC AGGGGGGTTG	CCTACCACCA GGATGGTGGT
4951	CCACCACCAC GGTGGTGGTG	CACCTCCTCC GTGGAGGAGG	CCCCTCGCTG GGGGAGCGAC	CCGGACGACG GGCCTGCTGC	AGCTCCTCCC TCGAGGAGGG
5001	CCCTCCCCCT GGGAGGGGGA	CCGCCGCCGC GGCGGCGGCG	CGGTAACCAC GCCATTGGTG	CCCGCCCCCTC GGGCGGGGAG	TCCTCTTTCT AGGAGAAAGA

FIG._40G

82 / 154

5051 TTCTCCGTTT TTTTTTTCGT CTCGGTCTCG ATCTTTGGCC TTGGTAGTTT
AAGAGGCAAA AAAAAAAGCA GAGCCAGAGC TAGAAACCGG AACCATCAAA

5101 GGGTGGGCGA GAGCGGCTTC GTCGCCCAGA TCGGTGCGCG GGAGGGGCGG
CCCACCCGCT CTCGCCGAAG CAGCGGGTCT AGCCACGCGC CCTCCCCGCC

BamHI

~~~~~

5151 GATCTCGCGG CTGGCGTCTC CGGGCGTGAG TCGGCCCGGA TCCTCGCGGG  
CTAGAGCGCC GACCGCAGAG GCCCGCACTC AGCCGGGCCT AGGAGCGCCC

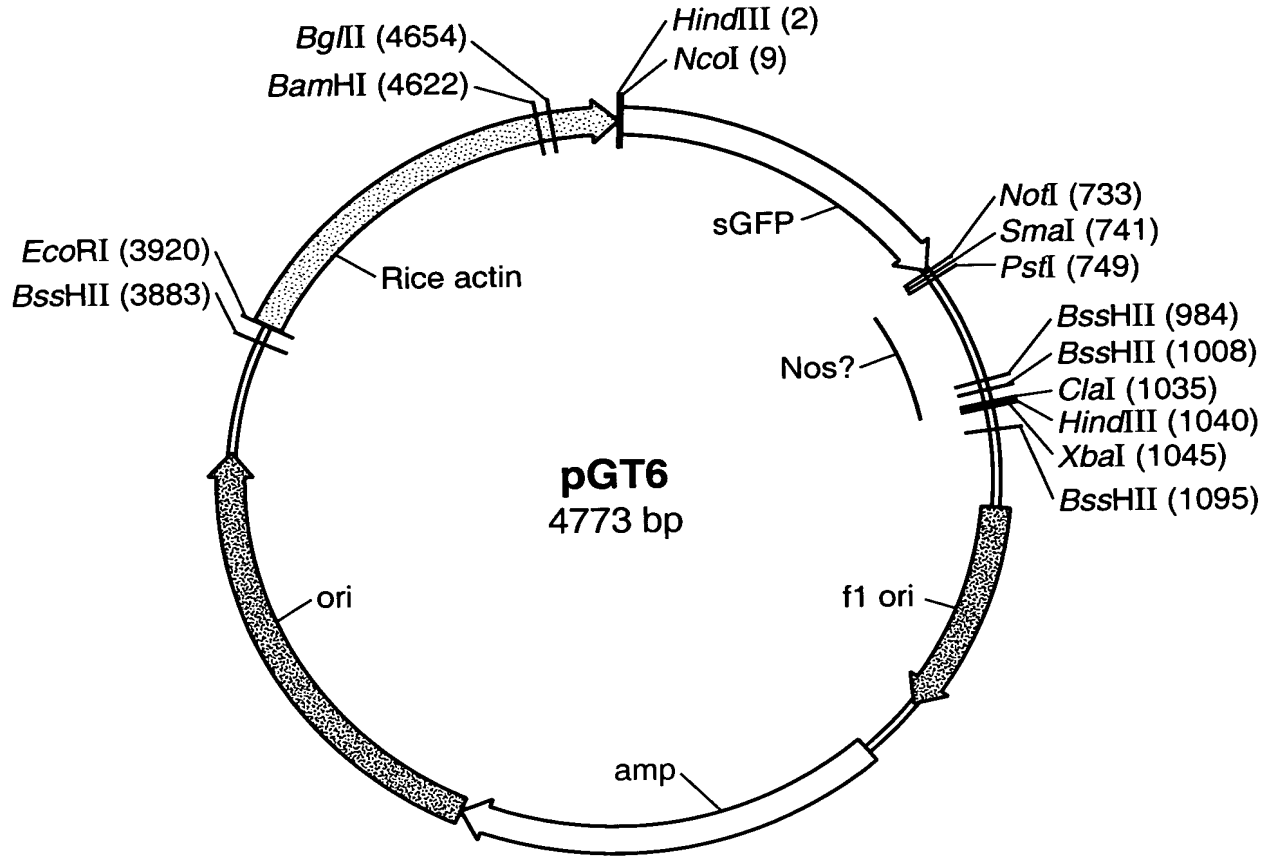
5201 GAATGGGGCT CTCGGATGTA GATCTTCTTT CTTTCTTCTT TTTGTGGTAG  
CTTACCCCGA GAGCCTACAT CTAGAAGAAA GAAAGAAGAA AACACCATC

5251 AATTTGAATC CCTCAGCATT GTTCATCGGT AGTTTTTCTT TTCATGATTT  
TTAAACTTAG GGAGTCGTAA CAAGTAGCCA TCAAAAAGAA AAGTACTAAA

5301 GTGACAAATG CAGCCTCGTG CGGAGCTTTT TTGTAGC  
CACTGTTTAC GTCGGAGCAC GCCTCGAAAA AACATCG

**FIG.\_40H**

83 / 154



**FIG.\_41A**

+

HindIII NcoI

~~~~~

1 AAGCTTACCA TGGTGAGCAA GGGCGAGGAG CTGTTACCG GGGTGGTGCC CATCTGGTC GAGCTGGACG
TTCGAATGGT ACCACTCGTT CCCGCTCCTC GACAAAGTGGC CCCACCACGG GTAGGACCAG CTCGACCTGC

71 GCGACGTGAA CGGCCACAAG TTCAGCGTGT CCGGCGAGGG CGAGGGCGAT GCCACCTACG GCAAGCTGAC
CGCTGCACTT GCCGGTGTTC AAGTCGCACA GCGCGCTCCC GCTCCCGCTA CCGTGGATGC CGTTCGACTG

141 CCTGAAGTTC ATCTGCACCA CCGGCAAGCT GCCCGTGCCC TGGCCACACC TCGTGACCAC CTTACACCTAC
GGACTTCAAG TAGACGTGGT GGCCGTTCGA CCGGCACGGG ACCGGGTGG AGCACTGGTG GAAGTGGATG

211 GCGGTGCAGT GCTTCAGCCG CTACCCCGAC CACATGAAGC AGCACGACTT CTTCAAGTCC GCCATGCCCG
CCGCACGTCA CGAAGTCGGC GATGGGGCTG GTGTACTTCG TCGTGTGAA GAAGTTCAGG CCGTACGGGC

281 AAGGCTACGT CCAGGAGCGC ACCATCTTCT TCAAGGACGA CGGCAACTAC AAGACCCCGC CCGAGGTGAA
TTCCGATGCA GGTCCCTCGC TGGTAGAAGA AGTTCCTGCT GCCGTTGATG TTCTGGGCGC GGCTCCACTT

351 GTTCGAGGGC GACACCCCTGG TGAACCGCAT CGAGCTGAAG GGCATCGACT TCAAGGAGGA CGGCAACATC
CAAGCTCCCG CTGTGGGACC ACTTGGCGTA GCTCGACTTC CCGTAGCTGA AGTTCCTCCT GCCGTGTAG

421 CTGGGGCACA AGCTGGAGTA CAACTACAAC AGCCACAACG TCTATATCAT GGCCGACAAG CAGAAGAACG
GACCCCGTGT TCGACCTCAT GTTGATGTTG TCGGTGTGTC AGATATAGTA CCGGCTGTTT GTCTTCTTGC

491 GCATCAAGGT GAACTTCAAG ATCCGCCACA ACATCGAGGA CGGCAGCGTG CAGCTCGCCG ACCACTACCA
CGTAGTTCCA CTTGAAGTTC TAGGCGGTGT TGTAGCTCCT GCCGTCGCAC GTCGAGCGGC TGGTGATGGT

561 GCAGAACACC CCCATCGGGC ACGGCCCCGT GCTGCTGCCC GACAACCACT ACCTGAGCAC CCAGTCCGCC
CGTCTTGTGG GGGTAGCCGC TGCCGGGGCA CGACGACGGG CTGTTGGTGA TGGACTCGTG GGTACGGCGG

631 CTGAGCAAAG ACCCCAACGA GAAGCGCGAT CACATGGTCC TGCTGGAGTT CGTGACCGCC GCCGGGATCA
GACTCGTTTC TGGGGTTGCT CTTCCGCGCTA GTGTACCAGG ACGACCTCAA GCACTGGCGG CGGCCCTAGT

FIG..41B

+

85 / 154

```

SmaI
~~~~~
          NotI          PstI
          ~~~~~
701  CTCACGGCAT  GGACGAGCTG  TACAAAGTAA  GCGGCCGCC  GGGCTGCAGG  GAAACCACTG  AAGGATGAGC
    GAGTGCCGTA  CCTGCTCGAC  ATGTTCAATT  CGCCGGCGGG  CCGACGTCCT  CTTTGGTGAC  TTCTTACTCG

771  TGTAAGAAG  CAGATCGTTC  AAACATTTGG  CAATAAAGTT  TCTTAAGATT  GAATCCTGTT  GCCGGTCTTG
    ACATTTCTTC  GTCTAGCAAG  TTTGTAAACC  GTTATTTCAA  AGAATTCTAA  CTTAGGACAA  CGGCCAGAAC

841  CGATGATTAT  CATATAAATT  CTGTTGAATT  ACGTTAAGCA  TGTAATAAAT  AACATGTAAT  GCATGACGTT
    GCTACTAATA  GTATATTAA  GACAACTTAA  TGCAATTCGT  ACATTAATTA  TTGTACATTA  CGTACTGCAA

911  ATTTATGAGA  TGGGTTTTTA  TGATTAGAGT  CCGGCAATTA  TACATTTAAT  ACGCGATAGA  AAACAAATA
    TAAATACTCT  ACCCAAAAT  ACTAATCTCA  GGGCGTTAAT  ATGTAAATTA  TCGCGTATCT  TTTGTTTTAT

          XbaI
          ~~~~~

          ClaI HindIII
          ~~~~~
981  TAGCGCGCAA  ACTAGGATAA  ATTATCGCGC  GCGGTGTCTAT  CTATGTTACT  AGATCGATAA  GTTCTTAGAG
    ATCGCGCGTT  TGATCCTATT  TAATAGCGCG  CGCCACAGTA  GATACAATGA  TCTAGCTATT  CGAAGATCTC

          BssHII
          ~~~~~

          BssHII
          ~~~~~
1051  CGGCCGGTGG  AGCTCCAATT  CGCCCTATAG  TGAGTCGTAT  TACGCGCGCT  CACTGGCCGT  CGTTTACAA
    GCCGGCCACC  TCGAGGTTAA  GCGGGATATC  ACTCAGCATA  ATGCGCGCGA  GTGACCGGCA  GCAAAATGTT

1121  CGTCGTGACT  GGGAAAACCC  TGGCGTTACC  CAACTTAATC  GCCTTGCAGC  ACATCCCCCT  TTCGCCAGCT
    GCAGCACTGA  CCTTTTGGG  ACCGCAATGG  GTTGAATTAG  CGGAACGTCG  TGTAGGGGGA  AAGCGGTCGA

1191  GCGGTAATAG  CGAAGAGGCC  CGCACCGATC  GCCCTTCCCA  ACAGTTGCGC  AGCCTGAATG  GCGAATGGGA
    CCGCATTATC  GCTTCTCCGG  GCGTGGCTAG  CGGGAAGGGT  TGTCAACGCG  TCGGACTTAC  CGCTTACCCT

1261  CGCGCCCTGT  AGCGGCGCAT  TAAGCGCGGC  GGGTGTGGTG  GTTACGCGCA  GCGTGACCGC  TACACTTGCC
    GCGCGGGACA  TCGCCCGCGTA  ATTCGCGCCG  CCCACACCAC  CAATGCGCGT  CGCACTGGCG  ATGTGAACGG

```

FIG._41C

1331 AGCGCCCTAG CGCCCGCTCC TTTTCGCTTTC TTCCCTTCTT TTCTCGCCAC GTTCGCCGGC TTTCCCCGTC
TCGCGGGATC GCGGGCGAGG AAGCGAAAG AAGGAAGGA AAGAGCGGTG CAAGCGGCCG AAAGGGGCAG

1401 AAGCTCTAAA TCGGGGGCTC CCTTTAGGGT TCCGATTTAG TGCTTTACGG CACCTCGACC CCAAAAAACT
TTTCGAGATTT AGCCCCCGAG GGAATCCCA AGGCTAAATC ACGAAATGCC GTGGAGCTGG GGTTTTGTGA

1471 TGATTAGGGT GATGGTTTAC GTAGTGGCC ATCGCCCTGA TAGACGGTTT TTCGCCCTTT GACGTTGGAG
ACTAATCCCA CTACCAAGTG CATCACCCGG TAGCGGACT ATCTGCCAAA AAGCGGGAAG CTGCAACCTC

1541 TCCACGTTCT TTAATAGTGG ACTCTTGTTC CAAACTGGAA CAACACTCAA CCTATCTCG GTCTATCTTT
AGGTGCAAGA AATTATCACC TGAGAACAAAG GTTTGACCTT GTTGTGAGTT GGGATAGAGC CAGATAAGAA

1611 TTGATTTTATA AGGGATTTTG CCGATTTTCG CCTATTGGTT AAAAAATGAG CTGATTTAAC AAAAAATTTAA
AACTAAATAT TCCCTAAAC GGCTAAAGCC GGATAACCAA TTTTITTACTC GACTAAATG TTTTITTAAT

1681 CGCGAATTTT AACAAAAATAT TAACGCTTAC AATTAGGTG GCACTTTTCG GGGAAATGTG CGCGGAACCC
GCGCTTAAAA TTGTTTTTATA ATTGCGAATG TTAAATCCAC CGTGAAAAAGC CCCTTTACAC GCGCCTTGGG

1751 CTATTTGTTT ATTTTCTTAA ATACATTCAA ATATGTATCC GTCATGAGA CAATAACCCCT GATAAATGCT
GATAAACAAA TAAAAAGATT TATGTAAGTT TATACATAGG CGAGTACTCT GTTATTGGGA CTATTACGA

1821 TCAATAATAT TGA AAAAAGGA AGAGTATGAG TATTTCAACAT TTCCGTTGTCG CCCTTATTCC CTTTTTTGCG
AGTTATTATA ACTTTTTCCT TCTCATATCT ATAAAGTTGA AAGGCACAGC GGGAAATAAGG GAAAAAACGC

1891 GCATTTTGCC TTCTGTGTTT TGCTCACCCA GAACGCTGG TGAAGTAAA AGATGCTGAA GATCAGTTGG
CGTAAACCGG AAGGACAAA ACGAGTGGGT CTTTGCAGCC ACTTTTCATTT TCTACGACTT CTAGTCAACC

1961 GTGCACGAGT GGGTTACATC GAACTGGATC TCAACAGCGG TAAAGTCTTT GAGAGTTTTC GCCCGAAGA
CACGTGCTCA CCCAATGTAG CTTGACCTAG AGTTGTGCGC ATTCTAGGAA CTCTCAAAAG CCGGGCTTCT

2031 ACGTTTTCCA ATGATGAGCA CTTTTAAAGT TCTGCTATGT GCGCGGGTAT TATCCCGTAT TGACGCCGGG
TGCAAAAAGGT TACTACTCGT GAAAATTTC AAGACGATACA CCGCGCCATA ATAGGGCATA ACTGCGGGCC

2101 CAAGAGCAAC TCGGTCGCCG CATACACTAT TCTCAGAATG ACTTGGTTGA GTACTCACCA GTCAACAGAAA
GTTCTCGTTG AGCCAGCGGC GTATGTGATA AGAGTCTTAC TGAACCAACT CATGAGTGGT CAGTGTCTTT

FIG.-41D

87 / 154

2171 AGCATCTTAC GGATGGCATG ACAGTAAGAG AATTATGCAG TGCTGCCATA ACCATGAGTG ATAACACTGC
 TCGTAGAATG CCTACCGTAC TGTCATTCTC TTAATACGTC ACGACGGTAT TGGTACTCAC TATTGTGACG

2241 GGCCAACTTA CTTCTGACAA CGATCGGAGG ACCGAAGGAG CTAACCGCTT TTTTGCACAA CATGGGGGAT
 CCGGTTGAAT GAAGACTGTT GCTAGCCTCC TGGCTTCCTC GATTGGCGAA AAAACGTGTT GTACCCCTTA

2311 CATGTAACTC GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC AAACGACGAG CGTGACACCA
 GTACATTGAG CGGAAC TAGC AACCCTTGGC CTCGACTTAC TTCGGTATGG TTTGCTGCTC GCACTGTGGT

2381 CGATGCCTGT AGCAATGGCA ACAACGTTGC GCAAACCTATT AACTGGCGAA CTACTTACTC TAGCTTCCCG
 GCTACGGACA TCGTTACCGT TGTGTCAACG CGTTTGATAA TTGACCGCTT GATGAATGAG ATCGAAGGGC

2451 GCAACAATTA ATAGACTGGA TGGAGGCGGA TAAAGTTGCA GGACCACCTC TGCCTCGGC CCTTCCGGCT
 CGTTGTTAAT TATCTGACCT ACCTCCGCTT ATTTCAACGT CCTGGTGAAG ACGCGAGCCG GGAAGGCCGA

2521 GGCTGGTTTA TTGCTGATAA ATCTGGAGCC GGTGAGCGTG GGTCTCGCGG TATCATTTGCA GCACTGGGGC
 CCGACCAAAT AACGACTATT TAGACCTCGG CCACTCGCAC CCAGAGCGCC ATAGTAACGT CGTGACCCCG

2591 CAGATGGTAA GCCCTCCCGT ATCGTAGTTA TCTACACGAC GGGGAGTCAG GCAACTATGG ATGAACGAAA
 GTCTACCATT CCGGAGGGCA TAGCATCAAT AGATGTGCTG CCCCTCAGTC CGTTGATACC TACTTGCTTT

2661 TAGACAGATC GCTGAGATAG GTGCCCTCACT GATTAAGCAT TGGTAACTGT CAGACCAAGT TTACTCATAT
 ATCTGTCTAG CGACTCTATC CACGGAGTGA CTAATTCTGTA ACCATTGACA GTCTGGTTCA AATGAGTATA

2731 ATACTTTAGA TTGATTTAAA ACTTCATTTT TAATTTTAAA GGATCTAGGT GAAGATCCCT TTTGATAAATC
 TATGAAATCT AACTAAATTT TGAAGTAAAA ATTAATTTT CCTAGATCCA CTTCTAGGAA AAACATATTAG

2801 TCATGACCAA AATCCCTTAA CGTGAGTTTT CGTCCCACTG AGCGTCAGAC CCCGTAGAAA AGATCAAAGG
 AGTACTGGTT TTAGGGAATT GCACTCAAAA GCAAGGTGAC TCGCAGTCTG GGGCATCTTT TCTAGTTTCC

2871 ATCTTCTTGA GATCCTTTTT TTCTGCGCGT AATCTGCTGC TTGCAAAACAA AAAAACCCACC GCTACCAGCG
 TAGAAGAACT CTAGGAAAAA AAGACGGCA TTAGACGACG AACGTTTGTT TTTTGTGGTG CGATGGTCCG

2941 GTGGTTTGGT TGCCGGATCA AGAGCTACCA ACTCTTTTTC CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA
 CACCAACAAA ACGGCCCTAGT TCTCGATGGT TGAGAAAAAG GCTTCCATTG ACCGAAGTCG TCTCGCGTCT

FIG. 41E

88 / 154

```

3011 TACCAAATAC TGTCCCTTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG CACCGCCTAC
    ATGGTTTATG ACAGGAAGAT CACATCGGCA TCAATCCGGT GGTGAAGTTC TTGAGACATC GTGGCGGATG

3081 ATACCTCGCT CTGCTAATCC TGTTACCAGT GGCTGCTGCC AGTGGCGATA AGTCGTGTCT TACCGGGTTG
    TATGGAGCGA GACGATTAGG ACAATGGTCA CCGACGACGG TCACCGCTAT TCAGCACAGA ATGGCCCAAC

3151 GACTCAAGAC GATAGTTACC GGATAAGCG GCGGGTCCG GCTGAACGGG GGGTTCGTGC ACACAGCCCA
    CTGAGTCTCG CTATCAATGG CCTATTCCGC GTGCCAGCC CGACTTGCCC CCCAAGCACG TGTGTCGGGT

3221 GCTTGGAGCG AACGACCTAC ACCGAACTGA GATACCTACA GCGTGAGCTA TGAGAAAGCG CCACGGCTTC
    CGAACCTCGC TTGCTGGATG TGGCTTGAAT TATGGATGT CGCACTCGAT ACTCTTTCGC GGTGCGAAGG

3291 CGAAGGGAGA AAGCGGACA GGTATCCGGT AAGCGGCAGG GTCGGAACAG GAGAGCGCAC GAGGGAGCTT
    GCTTCCCTCT TTCCGCCTGT CCATAGGCCA TTCGCCGTCC CAGCCTTGTC CTCTCGCGTG CTCCCTCGAA

3361 CCAGGGGGAA ACGCCTGGTA TCTTTATAGT CCTGTCCGGT TTCGCCACCT CTGACTTGAG CGTCGATTTT
    GGTCCCCCTT TGCGGACCAT AGAATATCA GGACAGCCCA AAGCGGTGA GACTGAACCT GCAGCTAAAA

3431 TGTGATGCTC GTCAGGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACGCG GCCTTTTAC GGTTCCTGGC
    ACACTACGAG CAGTCCCCC GCCTCGGATA CCTTTTGCG GTCGTTGCG CGGAAAAATG CCAAGGACCG

3501 CTTTTGCTGG CCTTTTGCTC ACATGTTCTT TCCTGCGTTA TCCCTGATT CTGTGGATAA CCGTATTACC
    GAAAAACGACC GGAAAAACGAG TGTACAAGAA AGGACGCAAT AGGGACTAA GACACCTATT GGCATAATGG

3571 GCCTTTGAGT GAGCTGATAC CGCTCGCCGC AGCCGAACGA CCGAGCGCAG CGAGTCAGTG AGCGAGGAAG
    CGGAAACTCA CTCGACTATG GCGAGCGGCG TCGGCTTGCT GGCTCGCGTC GCTCAGTCAC TCGCTCCTTC

3641 CGGAAGAGCG CCCAATACGC AAACCGCCTC TCCCGCGCG TTGGCCGATT CATTAATGCA GCTGGCACGA
    GCCTTCTCGC GGGTTATGCG TTTGGCGGAG AGGGCGCGC AACCGGCTAA GTAATTACGT CGACCGTGCT

3711 CAGGTTTCCC GACTGGAAAG CGGGCAGTGA GCGCAACGCA ATTAATGTGA GTTAGCTCAC TCATTAGGCA
    GTCCAAAGGG CTGACCTTTC GCCCGTCACT CGCGTTGCGT TAATTACACT CAATCGAGTG AGTAATCCGT

3781 CCCCAGGCTT TACACTTTAT GCTTCCGGCT CGTATGTTGT GTGGAATTGT GAGCGGATAA CAATTTTACA
    GGGGTCCGAA ATGTGAAATA CGAAGGCCGA GCATACAACA CACCTTAACA CTCGCCCTATT GTTAAAGTGT

```

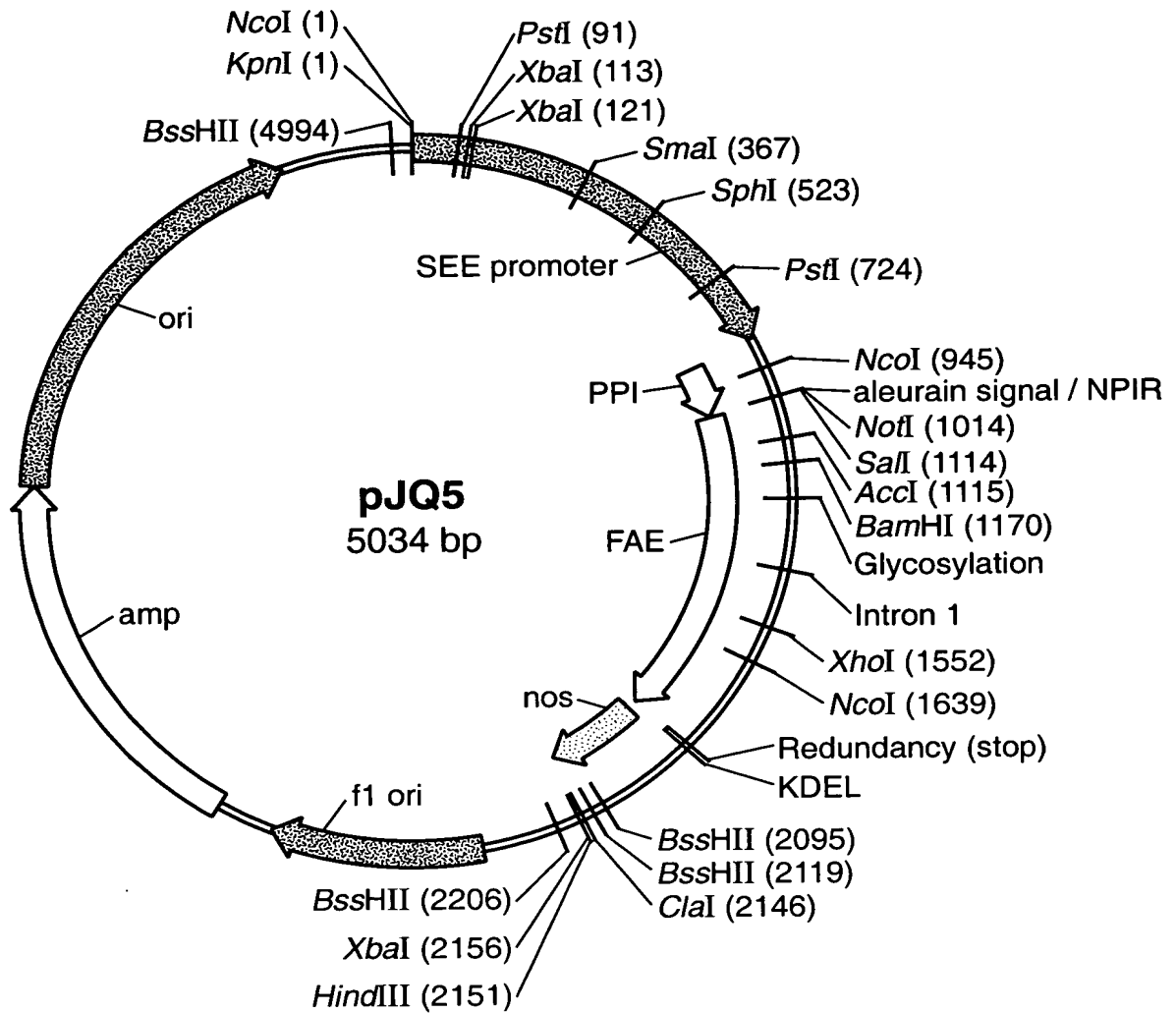
FIG._41F

FIG. 41G

BamHI BglII
~~~~~  
4621 GGATCCTCGC GGGGAATGGG GCTCTCGGAT GTAGATCTTC TTTCTTTCTTT CTTTTTGTGG TAGAATTGGA  
CCTAGGAGCG CCCCTTACCC CGAGAGCCCTA CATCTAGAAG AAAGAAAGAA GAAAAACACC ATCTTAAACT  
4691 ATCCCTCAGC ATTGTTTCATC GGTAAGTTTTT CTTTTCATGA TTTGTGACAA ATGCAGCCTC GTGCGGAGCT  
TAGGGAGTCG TAACAAGTAG CCATCAAAA GAAAGTACT AAACACTGTT TACGTCTGGAG CACGCCCTCGA  
4761 TTTTGTAGG TAG  
AAAAACATCC ATC

FIG.\_41H

91 / 154



**FIG. 42A**

92 / 154

```

NcoI
~~~~~
KpnI
~
1 CATGGGCCAG GTATAATTAT GGGATATCTC AAGCAAATAA TCGAAATATC ACCATTGGCT ACAATATCTG
 GTACCCGGTC CATATTAATA CCTATAGAG TTCGTTTATT AGCTTTATAG TGGTAACCGA TGTATATAGAC

 PstI XbaI XbaI
       ~~~~~                      ~~~~~
71  AGCTCCGAGT  TCTGACTGCA  GTC TGGATGA  CGCGTGTGTG  ATCTAGAACT  CTAGATAGCA  CAGCCACAGC
   TCGAGGCTCA  AGACTGACGT  CAGACCTACT  GCGCACAAAC  TAGATCTTGA  GATCTATCGT  GTCGGTGTCTG

141  ACCTACAGGA  GTGCGACACT  TGTGGACTGT  AGTAGTGTGT  GAGACGGAGC  TCTTTCCCTAC  CTCCTGACGT
   TGGATGTCTT  CACGCTGTGA  ACACCTGACA  TCATCACAAC  CTCTGCCTCG  AGAAAGGATG  GAGGACTGCA

211  TGCCGCCGTT  GTCCATTCCA  ACGGCATCAC  TCTCAACCAC  TCACGCGCTC  CCAACAAAAT  ATCGTCCCCC
   ACGCGGCCAA  CAGGTAAGGT  TGCCGTAGTG  AGAGTTGGTT  AGTGCGCGAG  GGTGTGTTTA  TAGCAGGGGG

281  ATGTCTTGGC  GGAGAGAGAG  TACATACATG  CTGTCGCGCC  GTTTTGTGCT  GAATCTCGCT  TCCACTGGCC
   TACAGAACCG  CCTCTCTCTC  ATGTATGTAC  GACAGCGCGG  CAAAACACA  CTTAGAGCGA  AGGTGACCGG

       SmaI
       ~~~~~
351 AATCAGCTCA GCTCCCGGGA GCTCACTCAT TCAAGATCCC ATCGTCGTCT TCACCCCCTGG CGTCATGGGA
 TTAGTCGAGT CGAGGGCCCT CGAGTGAGTA AGTTCTAGGG TAGCAGCAGC AGTGGGGACC GCAGTACCCT

421 TGGAAAAGAA CCTCCGTTGC TCGGATGAGT CAGCCATATC CCCGAACAGA GTACTGCAAG ATAAACCCAAT
 ACCTTTCTTT GGAGGCAACG AGCCTACTCA GTCGGTATAG GGGCTTGTCT CATGACGTTT TATTGGGTTA

 SphI
       ~~~~~
491  TCAGATTCCC  CCAATAGAGA  AAGTATAGCA  TGCTTTCGGG  TTTTGTGTTG  CTTAATTGAC  TTTATTTTGT
   AGTCTAAGGG  GGTATATCTT  TTCAATATCGT  ACGAAAGCCC  AAAACAAACC  GAATTAACTG  AAATAAAAC

561  TTGGAGTTGA  ATGCTGATTT  GTTGTGTAAA  ATGCCCAACC  ATCTGAATAT  CGAGACGGAT  AATAGGCTGG
   AACCTCAACT  TACGACTAAA  CAACACATTT  TACGGGTTGG  TAGACTTATA  GCTCTGCCTA  TTATCCGACC

```

FIG.\_42B

93 / 154

```

631  CTAATTAATT TATAGCAAGA TTCTGTAGTG CACATCGCAA ATATCTTTCT GGGCATTACA GCTGGAGGCT
    GATTAATTAA ATATCGTTCT AAGACATCAC GTGTAGCGTT TATAGAAAGA CCCGTAATGT CGACCTCCGA

                                PstI
                                ~~~~~
701 TCATCAGCCT GAAACACTCT GCAGAGCCTG AAGCAAGTGG TGAAGCGTGG CGATGAGATG GGTATAAAAC
 AGTAGTCGGA CTTTGTGAGA CGTCTCGGAC TTCGTTTACC ACTTCGCACC GCTACTCTAC CCATATTTTG

771 CCCCCGCACC GGGACGCGAG CTCCCGCCTA CCAGTACCAT CTCGCCCTGC TCCCCCTGCC GGACGACCCA
 GGGGCCGTGG CCTGCGCTC GAGGCGGAT GGTCAATGGTA GAGCGGAGCG AGGGGGACGG CCTGCTGGGT

841 GTAAAATACT GTTGCCCACT CGCCGGCGAG ATGGMCGTGC ACAAGGAGGT SAACCTTCGTS GCCTACCTCC
 CATTTATGA CAACGGGTGA CGGCCCGCTC TACCKGCACG TGTTCTCTCA STTGAAGCAS CGGATGGAGG

 NotI
                                ~~~~~
911  TGATCGTSCT CGGCCTCCTC TTGCTCGTST CCGCCATGGA GCACGTGGAC GCCAAGGCCT GCACCCKCGA
    ACTAGCASA GCCGGAGGAG AACGAGCASA GGCGGTACCT CGTGCACCTG CGGTTCCGGA CGTGGGMGCT

                                NotI
                                ~~~~~
981 GTGCGGCAAC CTCGGCTTCG GCATCTGCCC GGCGGCCGCC TCCACGCAGG GCATCTCCGA AGACCTCTAC
 CACGCCGTTG GAGCCGAAGC CGTAGACGGG CCGCCGGCGG AGGTGCGTCC CGTAGAGGCT TCTGGAGATG

 SalI
                                ~~~~~
                                AccI
                                ~~~~~
1051 AGCCGTTTAG TCGAAAATGGC CACTATCTCC CAAGTGCCTT ACGCCGACCT GTGCAACATT CCGTCGACTA
 TCGGCAAAATC AGCTTTACCG GTGATAGAGG GTTCGACGGA TCGGGCTGGA CACGTTGTAA GGCAGCTGAT

 BamHI
                                ~~~~~
1121 TTATCAAGGG AGAGAAAATT TACAATTCTC AAACGTGACAT TAACGGATGG ATCCTCCGCG ACGACAGCAG
    AATAGTTCCC TCTCTTTTAA ATGTTAAGAG TTTGACTGTA ATTGCCCTACC TAGGAGGCGC TGCTGTCTGC
  
```

FIG.. 42C

+

```
1191 CAAAGAAATA ATCACCCTCT TCCGTGGCAC TGGTAGTGAT ACGAATCTAC AACTCGATAC TAACTACACC
      GTTTCCTTTAT TAGTGGCAGA AGGCACCGTG ACCATCACTA TGCTTAGATG TTGAGCTATG ATTGATGTGG

1261 CTCACGCCCTT TCGACACCCCT ACCACAATGC AACGGTTGTG AAGTACACGG TGGATATTAT ATTGGATGGG
      GAGTGGGGAA AGCTGTGGGA TGGTGTACG TTGCCAACAC TTCATGTGCC ACCTATAATA TAACCTACCC

1331 TCTCCGTCCA GGACCAAGTC GAGTCGCTTG TCAAAACAGCA GGTTAGCCAG TATCCGGACT ACGCGCTGAC
      AGAGGCAGGT CCTGGTTCAG CTCAGCGAAC AGTTGTCTGT CCAATCGGTC ATAGGCCCTGA TCGCGGACTG

1401 CGTGACCGGC CACKCCCTCG GCGCCTCCCT GCGGGCACTC ACTGCCGCC AGCTGTCTGC GACATACGAC
      GCACTGGCCG GTGMGGGAGC CGCGGAGGGA CCGCCGTGAG TGACGGCGGG TCGACAGACG CTGTATGCTG

1471 AACATCCGCC TGTACACCTT CGGCGAACCG CGCAGCGGCA ATCAGGCCCTT CGCGTCGTAC ATGAACGATG
      TTGTAGGCGG ACATGTGGAA GCCGCTTGGC GCGTCGCCGT TAGTCCGGAA GCGCAGCATG TACTTGCTAC

      XhoI
      ~~~~~

1541 CCTTCCAAGC CTCGAGCCCA GATACGACGC AGTATTTCGG GGTCACTCAT GCCAACGACG GCATCCCAAA
 GGAAGGTTCT GAGCTCGGGT CTATGCTGCG TCATAAAGGC CCAGTGAGTA CCGTTGCTGC CGTAGGGTTT

 NcoI
      ~~~~~

1611 CCTGCCCCCG GTGGAGCAGG GGTACGCCCA TGGCGGTGTA GAGTACTGGA GCGTTGATCC TTACAGCGCC
      GGACGGGGGC CACCTCGTCC CCATGCGGGT ACCGCCACAT CTCATGACCT CGCAACTAGG AATGTCGCGG

1681 CAGAACACAT TTGTCTGCAC TGGGGATGAA GTGCAGTGCT GTGAGGCCCA CACTCCGGGT CCCGCTGTC
      GTCTTGTTGTA AACAGACGTG ACCCTACTTT CACGTCACGA CACTCCGGGT CCCGCTGTC CCACACTTAT

1751 ATGCGCACAC GACTTATTTT GGGATGACGA GCGGAGCCTG TACATGGTGA TCAGTCATTT CAGCCTCCCC
      TACGCGTGTG CTGAATAAAA CCCTACTGCT CGCCTCGGAC ATGTACCCT AGTCAGTAAA GTCGGAGGGG

1821 GAGTGTACCA GGAAAGATGG ATGTCCTGGA GAGGGGCGG CGTAACCACT GAAGGATGAG CTGTAAAGAA
      CTCACATGGT CCTTCTTACC TACAGGACCT CTCCCCCGGC GCATTGGTGA CTTCCCTACTC GACATTTCTT
```

FIG. 42D

+

+

95 / 154

```
1891 GCAGATCGTT CAAACATTTG GCAATAAAGT TTCTTAAGAT TGAATCCTGT TGCCGGTCTT GCGATGATTA
CGTCTAGCAA GTTTGTAAAC CGTTATTCTA AAGAAATTCTA ACTTAGGACA ACGGCCAGAA CGCTACTAAT

1961 TCATATAAAT TCTGTTGAAT TACGTTAAGC ATGTAATAAT TAACATGTAA TGCATGACGT TATTTATGAG
AGTATATTAA AGACAACCTA ATGCAATTCTG TACATTATTA ATTGTACATT ACGTACTGCA ATAAATACTC

2031 ATGGGTTTTT ATGATTAGAG TCCCGCAATT ATACATTTAA TACGCGATAG AAAACAATAA ATAGCGCGCA
TACCCAAAAA TACTAATCTC AGGCGGTAA TATGTAAATT ATGCGCTATC TTTTGTTTTA TATCGCGCGT

                BssHII
                ~~~~~
 ClaI HindIII
                ~~~~~
                XbaI
                ~~~~~

2101 AACTAGGATA AATTATCGCG CGCGGTGTCA TCTATGTTAC TAGATCGATA AGTTCTTAGA GCGGCCGGTG
TTGATCCTAT TTAATAGCGC GCGCCACAGT AGATACAATG ATCTAGCTAT TCGAAGATCT CGCCGGCCAC

 BssHII
                ~~~~~

2171 GAGCTCCAAT TCGCCCTATA GTGAGTCGTA TTACGCGCGC TCACTGGCCG TCGTTTTACA ACGTCGTGAC
CTCGAGGTTA AGCGGGATAT CACTCAGCAT AATGCGCGCG AGTGACCGGC AGCAAAATGT TGCAGCACTG

2241 TGGGAAAACC CTGGCGTTAC CCAACTTAAT CGCCTTGCG CACATCCCCC TTTCGCCAGC TGGCGTAATA
ACCCTTTTGG GACCGCAATG GGTGGAATTA GCGGAACGTC GTGTAGGGGG AAAGCGGTG ACCGCATTAT

2311 GCGAAGAGGC CCGCACCGAT CGCCCTTCCC AACAGTTGCG CAGCCTGAAT GGCGAATGGG ACGCGCCCTG
CGCTTCTCCG GCGGTGGCTA GCGGGAAGGG TTGTCAACGC GTCGGACTTA CCGCTTACCC TCGCGGGGAC

2381 TAGCGGCGCA TTAAGCGCGG CGGGTGTGGT GGTACGCGC AGCGTGACCG CTACACTTGC CAGCGCCCTA
ATCGCCGCGT AATTCGCGCC GCCCACACCA CCAATGCGCG TCGCACTGGC GATGTGAACG GTCGCGGGAT

2451 GCGCCCGCTC CTTTCGCTTT CTTCCCTTCC TTTCTCGCCA CGTTCCCGCG CTTTCCCGGT CAAGCTCTAA
CGCGGCGGAG GAAAGCGAAA GAAGGGAAGG AAAGAGCGGT GCAAGCGGCC GAAAGGGGCA GTTCGAGATT
```

FIG.\_42E

+

96 / 154

2521 ATCGGGGGCT CCTTTTAGGG TTCCGATTTA GTGCTTTACG GCACCTCGAC CCCAAAAAAC TTGATTAGGG  
TAGCCCCCGA GGGAAATCCC AAGGCTAAAT CACGAAATGC CGTGGAGCTG GGGTTTTTTG AACTAATCCC

2591 TGATGGTTCA CGTAGTGGG CATCGCCCTG ATAGACGGTT TTTCCGCCCTT TGACGTTGGA GTCCACGTTT  
ACTACCAAGT GCATCACCCG GTAGCGGGAC TATCTGCCAA AAGCGGGA ACTGCAACCT CAGGTGCAAG

2661 TTTAATAGTG GACTCTTGT CCAAACTGGA ACAACACTCA ACCCTATCTC GGCTATATCT TTTGATTATAT  
AAATTATCAC CTGAGAACAA GGTTTGACCT TGTTGTGAGT TGGGATAGAG CCAGATAAGA AACTAAATA

2731 AAGGGATTTT GCCGATTTG GCCTATTGGT TAAAAAATGA GCTGATTAA CAAAAATTTA ACGCCAATTT  
TTCCCTAAAA CGGCTAAAGC CGGATAACCA ATTTTACT CGACTAAAT GTTTTAAAT TCGCCTTAAA

2801 TAACAAAAATA TTAACGCTTA CAATTAGGT GGCACTTTTC GGGGAAATGT GCGCGGAACC CCTATTGTGT  
ATTGTTTAT AATTGCGAAT GTTAAATCCA CCGTGAAAAG CCCCTTTACA CCGCCCTTGG GGATAAACAA

2871 TATTTTCTA AATACATTCA AATATGTATC CGCTCATGAG ACAATAACCC TGATAAATGC TTCAATAATA  
ATAAAAAGAT TTATGTAAGT TTATACATAG GCGAGTACTC TGTATTGGG ACTATTACG AAGTTATTAT

2941 TTGAAAAAGG AAGAGTATGA GTATTCAACA TTTCCGTGTC GCCCTTATTC CCTTTTTCG GGCATTTTGC  
AACTTTTTC TTCTCATACT CATAAGTTGT AAAGGCACAG CCGGAATAAG GGAAAAAACG CCGTAAAAACG

3011 CTTCCCTGTT TTGCTCACCC AGAAACGCTG GTGAAAGTAA AAGATGCTGA AGATCAGTTG GGTGCACGAG  
GAAGGACAAA AACGAGTGGG TCTTTGCGAC CACTTTCATT TTCTACGACT TCTAGTCAAC CCACGTGCTC

3081 TGGGTTACAT CGAACTGGAT CTCAACAGCG GTAAGATCCT TGAGAGTTT TGCCCCGAAG AACGTTTTC  
ACCCAATGTA GCTTGACCCTA GAGTTGTCCG CATTCTAGGA ACTCTCAAA GCGGGGCTTC TTGCAAAAGG

3151 AATGATGAGC ACTTTTAAAG TTCTGCTATG TGGCGCGGTA TTATCCCGTA TTGACGCCGG GCAAGAGCAA  
TTACTACTCG TGAAAAATTC AAGACGATAC ACCGCGCCAT AATAGGGCAT AACTGCGGCC CGTTCTCGTT

3221 CTCGGTCGCC GCATACACTA TTCTCAGAAAT GACTTGGTTG AGTACTCACC AGTCACAGAA AAGCATCTTA  
GAGCCAGCGG CGTATGTGAT AAGAGTCTTA CTGAACCAAC TCATGAGTGG TCAGTGTCTT TTCGTAGAAT

3291 CGGATGGCAT GACAGTAAGA GAATTATGCA GTGCTGCCAT AACCATGAGT GATAACACTG CGGCCAACTT  
GCCTACCGTA CTGTCATTCT CTTAATACGT CACGACGGTA TTGGTACTCA CTATTGTGAC GCCGGTTGAA

FIG. 42F



97 / 154

3361 ACTTCTGACA ACGATCGGAG GACCGAAGGA GCTAACCGCT TTTTTCACACA ACATGGGGGA TCATGTAACT  
TGAAGACTGT TGCTAGCCTC CTGGCTTCCT CGATTGGCGA AAAAACGTGT TGTACCCCTT AGTACATTGA

3431 CGCCTTGATC GTTGGGAACC GGAGCTGAAT GAAGCCATAC CAAACGACGA GCGTGACACC ACGATGCCCTG  
GCGGAAC TAG CAACCCCTTG CCTCGACTTA CTTCGGTATG GTTTGCTGCT CGCACTGTGG TGCTACGGAC

3501 TAGCAATGGC AACAAAGTTG CGCAAACTAT TAAC TGCGGA ACTACTTACT CTAGCTTCCC GGCAACAATT  
ATCGTTACCG TTGTTGCAAC GCGTTTGATA ATTGACCGCT TGATGAATGA GATCGAAGGG CCGTTGTAA

3571 AATAGACTGG ATGGAGGCGG ATAAAGTTGC AGGACCACCTT CTGCGCTCGG CCCTTCCGGC TGGCTGGTTT  
TTATCTGACC TACCTCCGCC TATTTCAACG TCCTGTGAA GACGCGAGCC GGAAGGCCG ACCGACCAA

3641 ATTGCTGATA AATCTGGAGC CGGTGAGCGT GGGTCTCGCG GTATCATTCG AGCACTGGGG CCAGATGGTA  
TAACGACTAT TTAGACCTCG GCCACTCGCA CCCAGAGCGC CATAGTAACG TCGTGACCCC GGTCTACCAT

3711 AGCCCTCCCG TATCGTAGTT ATCTACACGA CGGGAGTCA GGCAACTATG GATGAACGAA ATAGACAGAT  
TCGGGAGGC ATAGCATCAA TAGATGTGCT GCCCTCAGT CCGTTGATAC CTACTTGCTT TATCTGTCTA

3781 CGCTGAGATA GGTGCCCTCAC TGATTAAGCA TTGGTAACTG TCAGACCAAG TTTACTCATA TATACTTTAG  
GCGACTCTAT CCACGGAGTG ACTAATTCTGT AACCATTCG AGTCTGGTTC AAATGAGTAT ATATGAAATC

3851 ATTGATTTAA AACTTCATTT TTAATTTAA AGGATCTAGG TGAAGATCCT TTTTGATAAT CTCTATGACCA  
TAACTAAATT TTGAAGTAAA AATTAAATTT TCCTAGATCC ACTTCTAGGA AAAACTATTA GAGTACTGGT

3921 AAATCCCTTA ACGTGAGTTT TCGTTCCACT GAGCGTCAGA CCCCGTAGAA AAGATCAAAG GATCTTCTTG  
TTTAGGGAAT TGCACCTCAA AGCAAGGTGA CTCGCAGTCT GGGGCATCTT TTCTAGTTTC CTAGAAGAAC

3991 AGATCCCTTT TTTCTGCGCG TAATCTGCTG CTTGCAAACA AAAAACCAC CGCTACCAGC GGTGGTTTGT  
TCTAGGAAAA AAAGACGCGC ATTAGACGAC GAACGTTTGT TTTTGTGGTG GCGATGGTTC CCACCAACA

4061 TTGCCGGATC AAGAGCTACC AACTCTTTT CCGAAGGTAA CTGGCTTCAG CAGAGCGCAG ATACCAATA  
AACGGCCTAG TTCTCGATGG TTGAGAAAAA GGCTTCCATT GACCGAAGTC GTCTCGCGTC TATGGTTTAT

4131 CTGTCTCTCT AGTGTAGCCG TAGTTAGGCC ACCACTTCAA GAACTCTGTA GCACCGCCTA CATACCTCGC  
GACAGGAAGA TCACATCGGC ATCAATCCGG TGGTGAAGTT CTTGAGACAT CGTGGCGGAT GTATGGAGCG

FIG.\_42G

98 / 154

4201 TCTGCTAATC CTGTTACCAG TGGCTGCTGC CAGTGGCGAT AAGTCGTGTC TTACCGGGTT GGA CTCAAGA  
AGACGATTAG GACAAATGGTC ACCGACGACG GTCACCGCTA TTCAGCACAG AATGGCCCAA CCTGAGTTCT

4271 CGATAGTTAC CGGATAAGGC GCAGCGGTGC GGCTGAACGG GGGGTTTCGTG CACACAGCCC AGCTTGGAGC  
GCTATCAATG GCCTATTCCG CGTCGCCAGC CCGACTTGCC CCCCAAGCAC GTGTGTCGGG TCGAACCTCG

4341 GAACGACCTA CACCGAACTG AGATACCTAC AGCGTAGCT ATGAGAAAGC GCCACGCTTC CCGAAGGGAG  
CTTGCTGGAT GTGGCTTGAC TCTATGGATG TCGCACTCGA TACTCTTTTCG CCGTGCGAAG GGCTTCCCTC

4411 AAAGCGGGAC AGGTATCCGG TAAAGCGCAG GGTCGGAACA GGAGAGCGCA CGAGGGAGCT TCCAGGGGGA  
TTTCCGCCCTG TCCATAGGCC ATTGCGCGTC CCAGCCTTGT CCTCTCGCGT GCTCCCTCGA AGTCCCCCT

4481 AACGCCTGGT ATCTTTATAG TCCTGTGCGG TTTCGCCACC TCTGACTTGA GCGTCGATTT TTGTGATGCT  
TTGCGGACCA TAGAAATATC AGGACAGCCC AAAGCGGTGG AGACTGAACT CGCAGCTAAA AACACTACGA

4551 CGTCAGGGG GCGGAGCCTA TGGAAAAACG CCAGCAACGC GGCCTTTTGA GCGTTCCTGG CCTTTTGTCTG  
GCAGTCCCCC CGCTTCGGAT ACCTTTGTG GGTGCTTGGC CCGGAAAAAT GCCAAGGACC GGAAAAACGAC

4621 GCCTTTTGCT CACATGTTCT TTCCCTGCGTT ATCCCCGTAT TCTGTGGATA ACCGTATTAC CGCCTTTGAG  
CGGAAAAACGA GTGTACAAGA AAGGACGCAA TAGGGGACTA AGACACCTAT TGGCATAATG GCGGAAACTC

4691 TGAGCTGATA CCGCTCGCCG CAGCCGAACG ACCGAGCGCA GCGAGTCAGT GAGCGAGGAA GCGGAAGAGC  
ACTCGACTAT GCGGAGCGGC GTCGGCTTGC TGGCTCGCGT CGCTCAGTCA CTCGCTCCTT CGCCTTCTCG

4761 GCCCAATACG CAAACCGCCT CTCCTCCGCGC GTTGGCCGAT TCATTAATGC AGCTGGCAGC ACAGGTTTCC  
CGGGTTATGC GTTTGGCGGA GAGGGCGCG CAACCGGCTA AGTAATTACG TCGACCGTGC TGTCCAAAGG

4831 CGACTGGAAA GCGGGCAGTG AGCGCAACGC AATTAATGTG AGTTAGCTCA CTCATTAGGC ACCCCAGGCT  
GCTGACCTTT CGCCCGTCAC TCGCGTTGCG TTAATTACAC TCAATCGAGT GAGTAATCCG TGGGGTCCGA

4901 TTACACTTTA TGCTTCCGGC TCGTATGTTG TGTGGAATTG TGAGCGGATA ACAATTTCAC ACAGGAAACA  
AATGTGAAAT ACGAAGGCCG AGCATACAAC ACACCTTAAC ACTCGCCTAT TGTAAAGTG TGTCTTTTGT

FIG.\_42H

09/991,209.070202

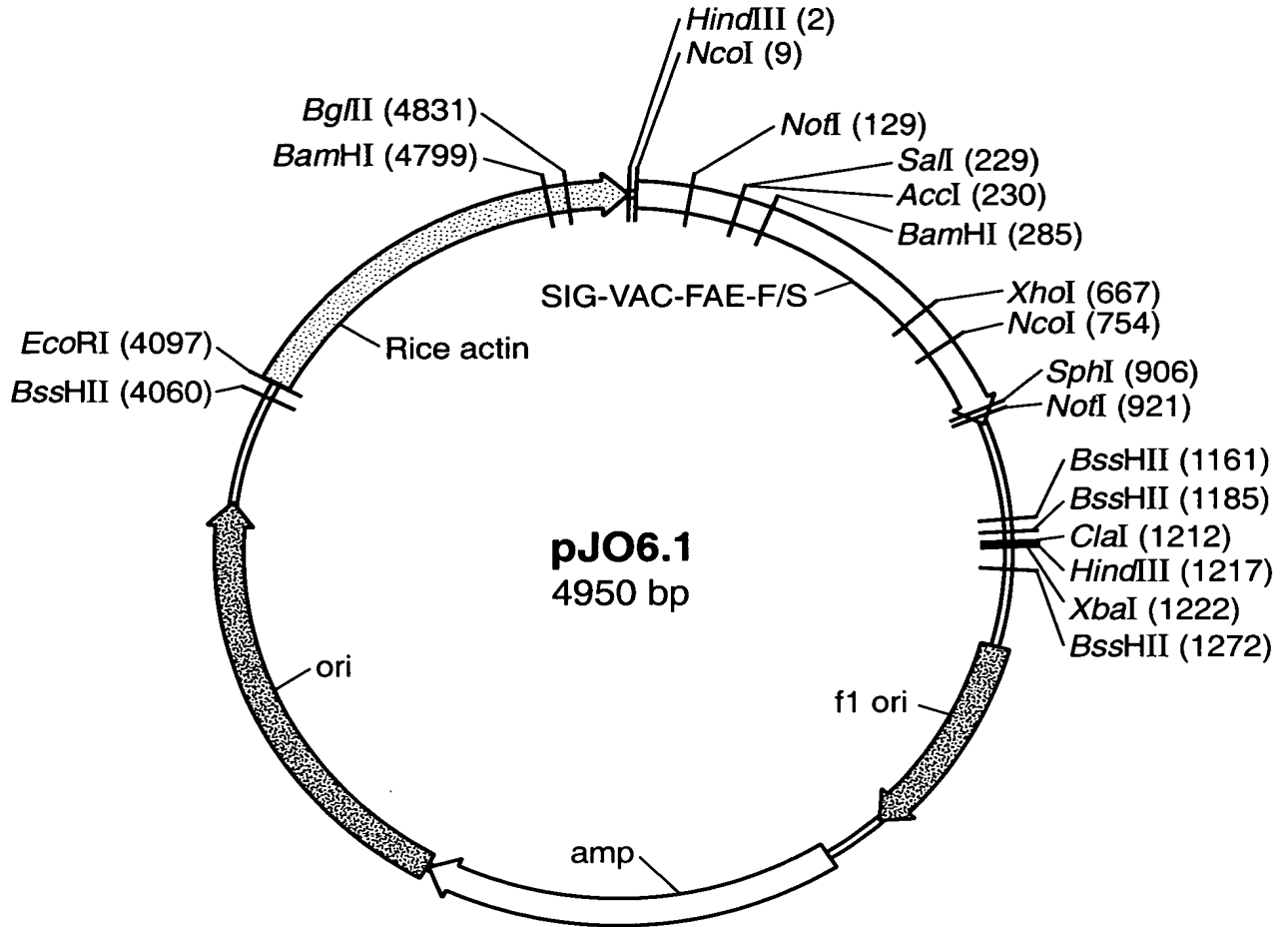
99 / 154

4971 GCTATGACCA TGATTACGCC AAGCGCGCAA TTAACCCCTCA CTAAAGGGAA CAAAAGCTGG GTAC  
CGATACTGGT ACTAATGCGG TTCGCGCGTT AATTGGGAGT GATTTCCTT GTTTTCGACC CATG

NcoI  
KpnI  
BssHII  
~~~~~  
~~~~~

FIG.\_42I

100 / 154



**FIG. 43A**

HindIII NcoI  
~~~~~  
1 AAGCTTACCA TGGCCACGC CCGCGTCCTC CTCCTGGCGC TCGCCGTGCT GCCACGGCC GCCGTCGCCG
TTCGAATGGT ACCGGGTGCG GCGCAGGAG GAGGACCGC GAGGCACGA CCGGTGCCG CGGCAGCGGC

NotI
~~~~~  
71 TCGCCTCCTC CTCCTCCTTC GCCGACTCCA ACCGATCCG GCCCGTCACC GACCGCGCG CCGCCTCCAC  
AGCGGAGGAG GAGGAGGAAG CCGCTGAGGT TGGCTAGGC CGGCAGTGG CTGGCGCGCC GCGGAGGTG

SalI  
~~~~~  
141 GCAGGGCATC TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCACTA TCTCCCAAGC TGCCTACGCC
CGTCCCGTAG AGGCTTCTGG AGATGTCGGC AAATCAGCTT TACCGGTGAT AGAGGTTCG ACGGATGCGG

AccI
~~~~~  
211 GACCTGTGCA ACATTCCGTC GACTATTATC AAGGGAGAGA AAATTACAA TTCTCAAACT GACATTAACG  
CTGGACACGT TGTAAAGGCG CTGATAATAG TTCCCTCTCT TTTAAATGTT AAGAGTTTGA CTGTAATTGC

BamHI  
~~~~~  
281 GATGGATCCT CCGCGACGAC AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCAC TGGTA GTGATACGAA
CTACCTAGGA GCGCTGCTG TCGTCGTTTC TTTATTAGTG GCAGAAGGCA CCGTGACCAT CACTATGCTT

351 TCTACAACAT GATACTAACT ACACCTCAC GCCTTTCGAC ACCCTACCCAC AATGCAACGG TTGTGAAGTA
AGATGTTGAG CTATGATTGA TGTGGGAGTG CCGAAAGCTG TGGGATGGTG TTACGTTGCC AACACTTCAT

421 CACGGTGGAT ATTATATTGG ATGGGTCTCC GTCCAGGACC AGTCGAGTC GCTTGTCAA CAGCAGGTTA
GTGCCACCTA TAATATAACC TACCCAGAGG CAGGTCCTGG TTCAGCTCAG CGAACAGTTT GTCGTCCAAT

491 GCCAGTATCC GGACTACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG CACTCAGTGC
CGGTACATAG CCTGATGCGC GACTGGCACT GGCCGGTGMG GGAGCCCGCG AGGGACCGCC GTGAGTGACG

561 CGCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC ACCTTCGGCG AACCGCGCAG CGGCAATCAG
GCGGTCGAC AGACGCTGTA TGCTGTTGTA GCGGACATG TGAAGCCGC TTGGCGCGTC GCCGTTAGTC

FIG._43B

102 / 154

| | | | |
|------|---|--------|--|
| | XhoI | | |
| 631 | GCCTTCGCGT CGTACATGAA CGATGCCTTC CAAGCCTCGA GCCCAGATAC GACGCAGTAT TTCCGGGTCA | | |
| | CGGAAGCGCA GCATGTACTT GCTACGGAAG GTTCGGAGCT CCGGTCTATG CTGCGTCATA AAGGCCCAGT | | |
| | | NcoI | |
| 701 | CTCATGCCAA CGACGGCATC CCAAACCTGC CCCCAGGTGA GCAGGGGTAC GCCCATGGCG GTGTAGAGTA | | |
| | GAGTACGGTT GCTGCCGTAG GGTTTGGACG GGGGCCACCT CGTCCCCATG CGGGTACCGC CACATCTCAT | | |
| 771 | CTGGAGCGTT GATCCTTACA GCGCCCAAGAA CACATTGTC TGCACCTGGG ATGAAGTGCA GTGCTGTGAG | | |
| | GACCTCGCAA CTAGGAATGT CGCGGTCTT GTGTAAACAG ACGTGACCCC TACTTCACGT CACGACACTC | | |
| | | SphI | |
| 841 | GCCCAGGGCG GACAGGGTGT GAATAATGCG CACACGACTT ATTTTGGGAT GACGAGCGGC GCATGCACCT | | |
| | CGGGTCCCGC CTGTCCCAACA CTATTACGC GTGTGCTGAA TAAACCCCTA CTGCTCGCCG CGTACGTGGA | | |
| | | NotI | |
| 911 | GGCCGGTTCG GCGCGCGGAA ACCACTGAAG GATGAGCTGT AAAGAAGCAG ATCGTTCAAA CATTGGCAA | | |
| | CCGGCCAGCG CCGCGCCCTT TGGTGACTTC CTACTCGACA TTTCTTCGTC TAGCAAGTTT GTAAACCGTT | | |
| 981 | TAAAGTTTCT TAAGATTGAA TCCCTGTTGCC GGTCTTGCGA TGATTATCAT ATAATTCTG TTGAATTACG | | |
| | ATTTCAAAGA ATTCTAACTT AGGACAAACGG CCAGAACGCT ACTAATAGTA TATTAAAGAC AACTTAATGC | | |
| 1051 | TTAAGCATGT AATAATTAC ATGTAATGCA TGACGTTATT TATGAGATGG GTTTTATGA TTAGAGTCCC | | |
| | AATCGTACA TTATTAAATG TACATTACGT ACTGCAATAA ATACTCTACC CAAAAATACT AATCTCAGG | | |
| | | BssHII | |
| 1121 | GCAATTATAC ATTTAATACG CGATAGAAA CAAAATATAG CGCGCAACT AGGATAAATT ATCGCGCGCG | | |
| | CGTTAATATG TAAATTATGC GCTATCTTTT GTTTTATATC GCGCGTTTGA TCCTATTATA TAGCGCGCGC | | |

FIG.. 43C

103 / 154

```

XbaI
~~~~~
ClaI HindIII
~~~~~
1191 GTGTCATCTA TGTTACTAGA TCGATAAGCT TCTAGAGCGG CCGGTGGAGC TCCAATTTCG CCTATAGTGA
CACAGTAGAT ACAATGATCT AGCTATTCTGA AGATCTCGCC GGCACCTCG AGGTTAAGCG GGATATCACT

BssHII
~~~~~
1261 GTCGTATTAC GCGCGCTCAC TGGCCGTCGT TTTACAAAGT CGTGACTGGG AAAACCCCTGG CGTTACCCAA
CAGCATAATG CCGCGGAGTG ACCGGCAGCA AAATGTTGCA GCACTGACCC TTTTGGGACC GCAATGGGTT

1331 CTTAATCGCC TTGCAGCACA TCCCCCTTTC GCCAGCTGGC GTAATAGCGA AGAGGCCCGC ACCGATCGCC
GAATTAGCGG AACGTCGTGT AGGGGAAAG CCGTCGACCG CATTATCGCT TCTCCGGGCG TGGCTAGCGG

1401 CTTCCCAACA GTTGGCAGC CTGAATGGG AATGGGACGC GCCCTGTAGC GCGCGCATTA GCGCGGCGGG
GAAAGGTTGT CAACGCGTCG GACTTACCGC TTACCCCTGCG CCGGACATCG CCGCGTAATT CGCGCCGCCC

1471 TGTGGTGGTT ACGCGCAGCG TGACCGCTAC ACTTGCCAGC GCCCTAGCGC CCGCTCCTTT CGCTTTCTTC
ACACCACCAA TCGCGGTCGC ACTGGCGATG TGAACGGTCG CCGGATCGCG GCGGAGGAAA GCGAAAGAAG

1541 CCTTCCTTTC TCGCCACGTT CGCCGGCTTT CCCCCTCAAG CTCTAAATCG GGGGCTCCCT TTAGGGTTCC
GGAAGGAAAG AGCGGTGCAA GCGGCCGAAA GGGGCAGTTC GAGATTTAGC CCCCAGGGA ATCCCAAGG

1611 GATTAGTGC TTTACGGCAC CTCGACCCCA AAAAATTGA TTAGGGTGAT GGTTCACGTA GTGGGCCATC
CTAAATCACG AAATGCCGTG GAGCTGGGGT TTTTGTGAAT ATCCCACTA CCAAGTGCAT CACCCGGTAG

1681 GCCCTGATAG ACGGTTTTC GCCCTTTGAC GTTGGAGTCC ACGTTCTTTA ATAGTGGACT CTTGTTCCAA
CGGGAATATC TGCCAAAAG CGGGAAACTG CAACCTCAGG TGCAAGAAAT TATCACCTGA GAACAAGGTT

1751 ACTGGAACAA CACTCAACCC TATCTCGGTC TATCTTTTG ATTTATAAGG GATTTGCGG ATTTCGGCCT
TGACCTTGTT GTGAGTTGGG ATAGAGCCAG ATAAGAAAAC TAAATATTCC CTAAAACGGC TAAAGCCGGA

1821 ATTGGTTAAA AAATGAGCTG ATTTAACAAA AATTAAACGC GAATTTTAAAC AAAATATTAA CGCTTACAAT
TAACCAATT TTTACTCGAC TAAATTGTTT TTAAATTGCG CTTAAATTTG TTTTATAAAT GCGAATGTTA

```

FIG._43D

104 / 154

```

1891  TTAGGTGGCA  CTTTTCGGGG  AAATGTGCGC  GGAACCCCTA  TTTGTTTATT  TTTCTAAATA  CATTCAAATA
      AATCCACCGT  GAAAAGCCCC  TTTACACGCG  CCTTGGGGAT  AAACAAATAA  AAAGATTAT  GTAAGTTTAT

1961  TGTATCCGCT  CATGAGACAA  TAACCCTGAT  AAATGCTTCA  ATAATATTGA  AAAAGGAAGA  GTATGAGTAT
      ACATAGGCCG  GTACTCTGTT  ATTGGGACTA  TTTACGAAGT  TATTATAACT  TTTTCCTTCT  CATACTCATA

2031  TCAACATTTC  CGTGTGCCCC  TTATTCCTTT  TTTTGGGGCA  TTTTGCCTTC  CTGTTTTTGC  TCACCCAGAA
      AGTTGTAAAG  GCACAGCGGG  AATAAGGGA  AAAACGCCGT  AAAACGGAAG  GACAAAACG  AGTGGGTCTT

2101  ACGCTGGTGA  AAGTAAAAGA  TGCTGAAGAT  CAGTTGGGTG  CACGAGTGGG  TTACATCGAA  CTGGATCTCA
      TGCAGACCACT  TTCATTTTCT  ACGACTTCTA  GTCAACCCAC  GTGCTCACCC  AATGTAGCTT  GACCTAGAGT

2171  ACAGCGGTAA  GATCCTTGAG  AGTTTTCGCC  CCGAAGAACG  TTTTCCAATG  ATGAGCACTT  TTAAAAGTTCT
      TGTGCGCCATT  CTAGGAACTC  TCAAAAGCGG  GGCTTCTTGC  AAAAGTTTAC  TACTCGTGAA  AATTTCAGA

2241  GCTATGTGGC  GCGGTATTAT  CCCGTATTGA  CGCCGGGGCA  GAGCAACTCG  GTCGCCCGCAT  ACACATTCTT
      CGATACACCG  CGCCATAATA  GGGCATAACT  CGGCCCCGTT  CTCGTTGAGC  CAGCGGCGTA  TGTGATAAGA

2311  CAGAATGACT  TGGTTGAGTA  CTCACCAGTC  ACAGAAAAGC  ATCTTACGGA  TGGCATGACA  GTAAGAGAAT
      GTCTTACTGA  ACCAACTCAT  GAGTGGTCAG  TGTCTTTTCG  TAGAATGCCCT  ACCGTACTGT  CATTCCTCTTA

2381  TATGCAGTGC  TGCCATAACC  ATGAGTGATA  ACACTGCGGC  CAACCTACTT  CTGACAAACGA  TCGGAGGACC
      ATACGTCACG  ACGGTATTGG  TACTCACTAT  TGTGACGCGG  GTTGAATGAA  GACTGTTGCT  AGCCTCCTGG

2451  GAAGGAGCTA  ACCGCTTTT  TGCACAACAT  GGGGATCAT  GTAACTCGCC  TTGATCGTTG  GGAACCGGAG
      CTTCTCTCGAT  TGGCGAAAAA  ACGTGTGTGA  CCCCCTAGTA  CATTGAGCGG  AACTAGCAAC  CCTTGGCCTC

2521  CTGAATGAAG  CCATACCAAA  CGACGAGCGT  GACACCACGA  TGCCTGTAGC  AATGGCAACA  ACGTTGCGCA
      GACTTACTTC  GGTATGGTTT  GCTGTCTCGA  CTGTGGTGCT  ACGGACATCG  TTACCCTTGT  TGCAACGCGT

2591  AACTATTAA  TGGCGAACTA  CTTACTCTAG  CTTCCCGGCA  ACAATTAAATA  GACTGGATGG  AGGCGGATAA
      TTGATAATTG  ACCGCTTGAT  GAATGAGATC  GAAGGGCCGT  TGTTAATTAT  CTGACCTACC  TCCGCCATT

2661  AGTTGCAGGA  CCACTTCTGC  GCTCGGCCCT  TCCGGCTGGC  TGTTTATTG  CTGATAAATC  TGGAGCCGGT
      TCAACGTCCT  GGTGAAGACG  CGAGCCGGGA  AGCCCGACCG  ACCAAATAAC  GACTATTAG  ACCTCGGCCA

```

FIG._43E

105 / 154

```

2731 GAGCGTGGGT CTCGCGGTAT CATTCAGCA CTGGGGCCAG ATGGTAAGCC CTCCCGTATC GTAGTTATCT
CTCGCACCCA GAGCGCCATA GTAACGTCGT GACCCCGGTC TACCATTCGG GAGGCGATAG CATCAATAGA

2801 ACACGACGGG GAGTCAGGCA ACTATGGATG AACGAAATAG ACAGATCGCT GAGATAGGTG CCTCACTGAT
TGTCGTGCCC CTCAGTCCGT TGATACCTAC TTGCTTTATC TGCTTAGCGA CTCTATCCAC GGAGTGAATA

2871 TAAGCATTTG TAACTGTCAG ACCAAGTTTA CTCATATATA CTTTAGATTG ATTTAAACT TCATTTTAA
ATTTCGTAACC ATGACAGTC TGGTTCAAAT GAGTATATAT GAAATCTAAC TAAATTTTGA AGTAAAAATT

2941 TTTAAAAGGA TCTAGGTGAA GATCCTTTT GATAATCTCA TGACCAAAAT CCTTAAACGT GAGTTTTCGT
AAATTTTCCT AGATCCACTT CTAGGAAAAA CTATTAGAGT ACTGGTTTAA GGAATTGCA CTCAAAAAGCA

3011 TCCACTGAGC GTCAGACCCC GTAGAAAAAG TCAAAAGGATC TTCTTGAGAT CCTTTTTC TGCGCGTAAT
AGGTGACTCG CAGTCTGGGG CATCTTTTCT AGTTTCCTAG AAGAACTCTA GAAAAAAAG ACGCGCATTA

3081 CTGCTGCTTG CAAACAAAAA AACCAACCGCT ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT
GACGACGAAC GTTTGTTTTT TTGGTGGCGA TGGTCGCCAC CAAACAAACG GCCTAGTTCT CGATGGTTGA

3151 CTTTTTCCGA AGGTAACCTG CTTCAGCAGA GCGCAGATAC CAAATACTGT CTTTCTAGTG TAGCCGTTAGT
GAAAAAGGCT TCCATTGACC GAAGTCGTCT CGCGTCTATG GTTTATGACA GGAAGATCAC ATCGGCATCA

3221 TAGGCCACCA CTTCAAGAAC TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCCTGT TACCAGTGGC
ATCCGGTGGT GAAGTTCTTG AGACATCGTG GCGGATGTAT GGAGCGAGAC GATTAGGACA ATGTCACCG

3291 TGCTGCCAGT GCGGATAAGT CGTGCTTTAC CGGTTGGAC TCAAGACGAT AGTTACCGGA TAAGGCGCAG
ACGACGGTCA CCGCTATTCA GCACAGAAATG GCCCAACCTG AGTTCTGCTA TCAATGGCCT ATTCCGCGTC

3361 CGGTCGGGCT GAACGGGGG TTCTGTGCACA CAGCCCAGCT TGGAGCGAAC GACCTACACC GAACGTAGAT
GCCAGCCCCG CTTGCCCCCC CAGCACGTGT GTCGGGTCGA ACCTCGCTTG CTGGATGTGG CTTGACTCTA

3431 ACCTACAGCG TGAGCTATGA GAAAGCGCCA CGCTTCCCGA AGGGAGAAAG GCGGACAGGT ATCCGGTAAG
TGGATGTCGC ACTCGATACT CTTTCGCGGT GCGAAGGGCT TCCCCTCTTC CGCCTGTCCA TAGGCCATTC

3501 CGGCAGGGTC GGAACAGGAG AGCGCACGAG GGAGCTTCCA GGGGAAACG CCTGGTATCT TTATAGTCCT
GCCGTCCTCG CCTTGTCCTC TCGCGTGCTC CCTCGAAGGT CCCCCTTTGC GGACCATAGA AATATCAGGA

```

FIG._43F

FIG. 43G

107 / 154

```
4271 ACCCAACCCA ACCACCCCA GTGCAGCCAA CTGGCAATA GTCTCCACCC CCGGCACATAT CACCGTGAGT
TGGGTGGGT TGGGTGGGT CACGTGGGT GACCGTTTAT CAGAGGTGG GCGCGTGATA GTGGCACTCA

4341 TGTCCGCACC ACCGCACGTC TCGCAGCCAA AAAAAAAAA AGAAGAAAA AAAAACAAGC
ACAGGCGTGG TGGCGTGCAG AGCGTCGGTT TTTTTTTTTT TCTTCTTTTT TTTTCTTTTT CTTTTTGTCTG

4411 AGGTGGGTCC GGTCTGTGGG GGCCGGAAAA GCGAGGAGGA TCGCGAGCAG CGACGAGGCC CGGCCCTCCC
TCCACCCAGG CCCAGCACC CCGGCCTTTT CGTCTCTCCT AGCGTCTGTC GCTGCTCCGG GCCGGGAGGG

4481 TCCGCTTCCA AAGAAACGCC CCCCATCGCC ACTATATACA TACCCCCCCC TCTCTCTCCA TCCCCCAAC
AGCGAAGGT TTCTTTGCGG GGGTAGCGG TGATATATGT ATGGGGGGG AGAGGAGGCT AGGGGGGTG

4551 CCTACCACCA CCACCAACC CACCTCCTCC CCCCTCGCTG CCGGACGACG AGCTCCTCCC CCCTCCCCCT
GGATGGTGGT GGTGGTGGTG GTGGAGGAGG GGGAGCGAC GGCCTGCTGC TCGAGGAGGG GGGAGGGGGA

4621 CCGCCGCCGC CGGTAACCAC CCCGCCCTC TCTCTTTCT TTCTCCTTTT TTTTCTTCTG CTCGGTCTCG
GGCGCGGCG GCCATTGGTG GGGCGGGAG AGGAGAAAGA AAGAGGCAAA AAAAAAGCA GAGCCAGAGC

4691 ATCTTTGGCC TTGGTAGTTT GGGTGGCGA GAGCGGCTC GTCGCCCAGA TCGGTGCGCG GGAGGGGCGG
TAGAAACCGG AACCATCAA CCCACCCGCT CTCGCCAAG CAGCGGTCT AGCCACGCGC CCTCCCCGCC

BglII BamHI
~~~~~
4761 GATCTCGCGG CTGGCGTCTC CGGGCGTGAG TCGGCCCGGA TCCTCGCGGG GAATGGGGCT CTCGGATGTA
CTAGAGCGCC GACCGCAGAG GCCCGCACTC AGCCGGGCT AGGAGCGCCC CTTACCCCGA GAGCCTACAT

BglII
~~~~~
4831 GATCTCTTTT CTTTCTTCTT TTTGTGGTAG AATTGAATC CCTCAGCATT GTTCATCGGT AGTTTTTCTT
CTAGAAGAAA GAAAGAAGAA AACACCATC TTAACCTTAG GGAGTCGTAA CAAGTAGCCA TCAAAAAAGAA

4901 TTTCATGATT GTGACAAATG CAGCCTCGTG CGGAGCTTTT TTGTAGGTAG
AAGTACTAAA CACTGTTTAC GTCGGAGCAC GCCTCGAAAA AACATCCATC
```

FIG..43H

108 / 154

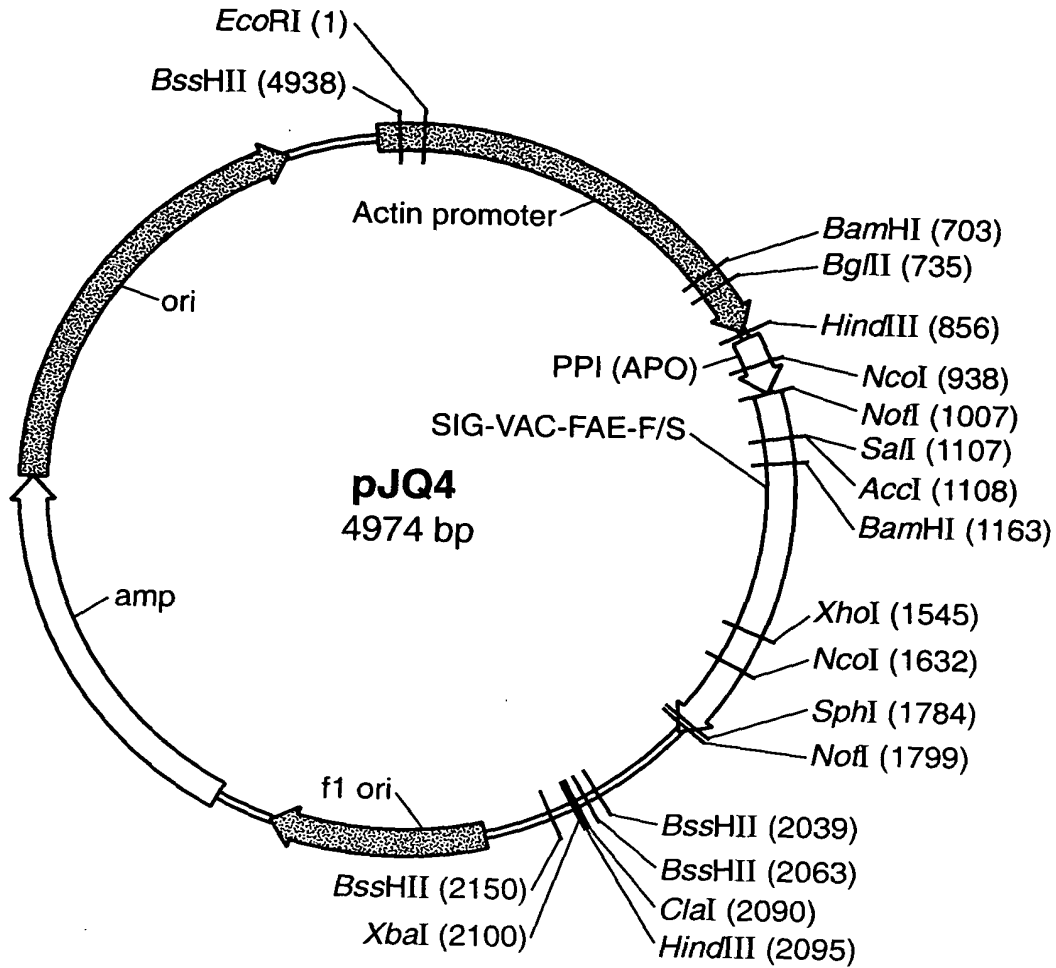


FIG. 44A

109 / 154

ECORI
~~~~~  
1 AATCCACAA TGAACAATAA TAAGATTAAA ATAGCTTGCC CCCGTTGCAG CGATGGGTAT TTTTCTTAGT  
TTAAGGTGTT ACTTGTTATT ATCTAATTT TATCGAACGG GGGCAACGTC GTACCCATA AAAAAGATCA  
  
71 AAAATAAAG ATAAACTTAG ACTCAAAAACA TTTACAAAAA CAACCCCTAA AGTCCTAAAG CCCAAAGTGC  
TTTTATTTC TATTGAATC TGAGTTTTGT AAATGTTTTT GTTGGGATT TCAGGATTTC GGGTTTCACG  
  
141 TATGCACGAT CCATAGCAAG CCCAGCCCAA CCCAACCCAC CCCAGTGCAG CCAACTGGCA  
ATACGTGCTA GGTATCGTTC GGGTCGGGT GGGTTGGGT GGGTGGGTG GGTTCACGTC GGTGACCGT  
  
211 AATAGTCTCC ACCCCCGGCA CTATCACCGT GAGTTGTCCG CACCACCGCA CGTCTCGCAG CCAAAAAAAA  
TTATCAGAGG TGGGGGCCGT GATAGTGCA CTCAACAGGC GTGGTGGCGT GCAGAGCGTC GGTTTTTTTT  
  
281 AAAAAGAAAG AAAAAGAAAG AAAAAGAAAG CAGCAGGTGG GTCCGGGTGG TGGGGGCCGG AAAAGCGAGG  
TTTTTCTTTC TTTTTTTTCT TTTTCTTTT GTCTCCACC CAGGCCCAGC ACCCCCGGCC TTTTTCGCTCC  
  
351 AGGATCGCGA GCAGCGACGA GGCCCCGCC TCCCTCCGCT TCCAAAGAAA CGCCCCCCCAT CGCCACTATA  
TCCTAGCGCT CGTCGCTGCT CCGGGCCGG AGGAGGCCA AGTTTCTTT GCGGGGGTA GCGGTGATAT  
  
421 TACATACCCC CCCCTCTCCT CCCATCCCC CAACCTACC ACCACCACCA CCACCACCTC CTCCCCCCTC  
ATGTATGGG GGGGAGAGGA GGTAGGGGG GTTGGGATGG TGGTGGTGGT GGTGGTGGAG GAGGGGGAG  
  
491 GCTGCCGGAC GACGAGCTCC TCCCCCTCC CCCCTCCCGG CGCCCGGTAA CCACCCCGCC CCTCTCCTCT  
CGACGGCCTG CTGCTCGAG AGGGGGGAG GGGAGCGGC GCGGGCCATT GGTGGGGCGG GGAGAGGAGA  
  
561 TTCTTTCTCC GTTTTTTTT TCGTCTCGGT CTCGATCTTT GGCTTTGGTA GTTTGGGTGG GCGAGAGCGG  
AAGAAAGAGG CAAAAAAGA AGCAGAGCCA GAGCTAGAAA CCGGAACCAT CAAACCCACC CGCTCTCGCC  
  
631 CTTCGTCGCC CAGATCGGTG CCGGGGAGGG GCGGGATCTC GCGGCTGGCG TCTCCGGGCG TGAGTCGGCC  
GAAGCAGCGG GTCTAGCCAC GCGCCCTCCC GCGCCCTAGAG CCGCGACCGC AGAGGCCCGC ACTCAGCCCG  
  
BamHI  
~~~~~  
701 CGGATCCTCG CGGGGAATGG GGCCTCTCGA TGATAGCTTT CTTCCTTTCT TCTTTTGTG GTAGAATTG
GCCTAGGAGC GCCCCCTTACC CCGAGAGCCT ACATCTAGAA GAAAGAAACA AGAAAAACAC CATCTTAAAC

FIG.. 44B

110 / 154

```

771  AATCCCTCAG CATTGTTTCAT CGGTAGTTTTT TCTTTTCATG ATTTGTGACA AATGCAGCCT CGTGCGGAGC
    TAGGGAGTC GTAACAAAGTA GCCATCAAAA AGAAAAGTAC TAAACACTGT TTACGTCGGA GCACGCCCTCG

    HindIII
    ~~~~~
841  TTTTTGTAG GTAGAAGCTT ACMATGCMCG TGCACAAGGA GGTSAACTTC GTSGCCTACC TCCTGATCGT
    AAAAAACATC CATCTTCGAA TGKTACCKGC ACGTGTTCCT CCASTTGAAG CASC GGATGG AGGACTAGCA

    NcoI
    ~~~~~
911  SCTCGGCTC CTCTTGCTCG TSTCCGCCAT GGAGCACGTG GACGCCAAGG CCTGCACCCK CGAGTGCGGC
    SGAGCCGGAG GAGAACGAGC ASAGGCGGTA CCTCGTGCAC CTGCGGTTCC GGACGTGGGM GCTCACGCCG

    NotI
    ~~~~~
981  AACCTCGGT TCGGCATCTG CCCGGCGGCC GCCTCCACGC AGGCATCTC CGAAGACCTC TACAGCCGTT
    TTGGAGCCGA AGCCGTAGAC GGGCCGCCGG CGGAGGTGCG TCCC GTAGAG GCTTCTGGAG ATGTCGGCAA

    SalI
    ~~~~~
    AccI
    ~~~~~
1051 TAGTCGAAAT GGCCACTATC TCCCAAGCTG CCTACGCCGA CCTGTGCAAC ATTCCGTCGA CTATTATCAA
    ATCAGCTTTA CCGGTGATAG AGGTTTCGAC GGATGCGGCT GGACACGTTG TAAGGCAGCT GATAATAGTT

    BamHI
    ~~~~~
1121 GGGAGAGAAA ATTTACAATT CTCAAACTGA CATTAACGGA TGGATCCCTC GCGACGACAG CAGCAAAGAA
    CCCTCTCTTT TAAATGTTAA GAGTTTGACT GTAATGCGCT ACCTAGGAGG CGCTGCTGTC GTCGTTTCTT

1191 ATAATCACCG TCTTCCGTGG CACTGGTAGT GATACGAATC TACAACCTCGA TACTAACTAC ACCCTCACGC
    TATTAGTGGC AGAAGGCACC GTGACCATCA CTATGCTTAG ATGTTGAGCT ATGATTGATG TGGGAGTGCG

1261 CTTTTCGACAC CCTACCACAA TGCAACGGTT GTGAAGTACA CGGTGGATAT TATATTGGAT GGTCTCCGT
    GAAAGCTGTG GGATGGTGTT ACGTTGCCAA CACTTCATGT GCCACCTATA ATATAACCTA CCCAGAGGCA
```

FIG.-44C

111 / 154

```

1331 CCAGGACCAA GTCGAGTCGC TTGTCAAACA GCAGGTTAGC CAGTATCCGG ACTACGCGCT GACCGTGACC
      GGTCCCTGGTT CAGCTCAGCG AACAGTTTGT CGTCCAATCG GTCATAGGCC TGATGCGCGA CTGGCACTGG

1401 GGCCACKCCC TCGGCGCCCTC CCTGGCGGCA CTCACGTGCCG CCCAGCTGTC TCGACATAC GACAACATCC
      CCGGTGMGGG AGCCGCGGAG GGACCGCCGT GAGTGACGGC GGTCTGACAG ACGCTGTATG CTGTTGTAGG

1471 GCCTGTACAC CTTCGGCGAA CCGCGCAGCG GCAATCAGGC CTTCCGCTCG TACATGAACG ATGCCCTTCCA
      CGGACATGTG GAAGCCGCTT GCGCGCTCGC CGTTAGTCCG GAAGCGCAGC ATGTACTTGC TACGGAAGGT

      XhoI
      ~~~~~
1541 AGCCTCGAGC CCAGATACGA CGCAGTATTT CCGGGTCACT CATGCCAACG ACGGCATCCC AAACCTGCCC
      TCGGAGCTCG GGTCTATGCT GCGTCATAAA GGCCCAGTGA GTACGGTTGC TGCCGTAGGG TTTGGACGGG

      NgoI
      ~~~~~
1611 CCGGTGGAGC AGGGGTACGC CCATGGCGGT GTAGAGTACT GGAGCGTTGA TCCTTACAGC GCCCAGAACA
      GGCCACCTCG TCCCCATGCG GGTACCGCCA CATCTCATGA CCTCGCAACT AGGAATGTCG CGGGTCTTGT

1681 CATTGTCTG CACTGGGGAT GAAGTGCAGT GCTGTGAGGC CCAGGGCGGA CAGGGTGTGA ATAATGCGCA
      GTAAACAGAC GTGACCCCTA CTTACAGTCA CGACACTCCG GGTCCCGCCT GTCCACACT TATTACGCGT

      SphI
      ~~~~~
      NotI
      ~~~~~
1751 CACGACTTAT TTTGGGATGA CGAGCGGCGC ATGCACCTGG CCGGTCGCGG CCGCGGAAC CACTGAAGGA
      GTGCTGAATA AAACCTACT GCTCGCCGCG TACGTGGACC GGCCAGCGCC GCGCCTTTG GTGACTTCCT

1821 TGAGCTGTAA AGAAGCAGAT CGTTCAAACA TTTGGCAATA AAGTTTCTTA AGATTGAATC CTGTTGCCGG
      ACTCGACATT TCTTCGTCTA GCAAGTTTGT AAACCGTTAT TTCAAAGAAT TCTAACTTAG GACAACGGCC

1891 TCTTGCATG ATTATCATAT AATTCTGTT GAATTACGTT AAGCATGTAA TAATTAACAT GTAATGCATG
      AGAACGCTAC TAATAGTATA TTAAGACAA CTTAATGCAA TTCGTACATT ATTAATTGTA CATTACGTAC

1961 ACGTTATTTA TGAGATGGGT TTTTATGATT AGAGTCCCGC AATTATACAT TTAATACGCG ATAGAAAACA
      TGAATAAAT ACTCTACCCA AAAATACTAA TCTCAGGGCG TTAATATGTA AATTATGCGC TATCTTTTGT

```

FIG._44D

| | BssHII
~~~~~ | BssHII
~~~~~ | BssHII
~~~~~ | Clal HindIII
~~~~~ | XbaI
~~ |
|------|---|-----------------|-----------------|-----------------------|------------|
| 2031 | AAATATAGCG CGCAAACTAG GATAAATTAT CGCGCGCGGT GTCATCTATG TTACTAGATC GATAAGCTTC TTTATATATCGC GCGTTTGATC CTATTTAATA GCGCGCGCCA CAGTAGATAC AATGATCTAG CTATTCGAAG | | | | |
| | XbaI
~~~~ | | BssHII
~~~~~ | | |
| 2101 | TAGAGCGGCC GGTGGAGCTC CAATTGCGCC TATAGTGAGT CGTATTACGC GCGCTCACTG GCCGTCGTTT ATCTCGCCGG CCACCTCGAG GTTAAGCGGG ATATCACTCA GCATAATGCG CGCGAGTGAC CGGCAGCAAA | | | | |
| 2171 | TACAACGTCTG TGA CTGGGAA AACCTGGCG TTAACCCAACT TAATCGCCCT GCAGCACATC CCCCTTTCGC ATGTTGCAGC ACTGACCCTT TTGGGACCGC AATGGGTTGA ATTAGCGGAA CGTCGTGTAG GGGGAAAGCG | | | | |
| 2241 | CAGCTGGCGT AATAGCGAAG AGGCCCGCAC CGATCGCCCT TCCCAACAGT TCGCGAGCCT GAATGGCGGAA GTCGACCGCA TTATCGCTTC TCCGGGCGTG GCTAGCGGGA AGGTTGTCA ACGCGTCGGA CTTACCGCTT | | | | |
| 2311 | TGGGACGCGC CCTGTAGCGG CGCATTAAGC GCGCGGGTG TGGTGGTTAC GCGCAGCGTG ACCGCTACAC ACCCTGCGG GGACATCGCC GCGTAATTCG CGCCGCCAC ACCACCAATG CGCGTCGCAC TGCGGATGTG | | | | |
| 2381 | TTGCCAGCGC CCTAGCGCC GCTCCTTTCG CTTTCTTCC TTCTTTTCTC GCCACGTTCTG CCGGCTTTC CCGGTCGCG GGATCGCGG CGAGGAAAGC GAAAGAAAG AGGTAAGAG CCGTGCAAGC GGCCGAAAGG | | | | |
| 2451 | CCGTCAAGCT CTAAATCGGG GGCTCCCTTT AGGGTTCCGA TTTAGTGCTT TACGGCACCT CGACCCCAA GGCAGTTCGA GATTAGCCC CCGAGGGAAA TCCCAAGGCT AAATCACGAA ATGCCGTGGA GCTGGGGTTT | | | | |
| 2521 | AAACTTGATT AGGGTGATGG TTCACGTA GTTACCGCATCA CCCGGTAGCG GGA CTATCTG CCTTTTTCG CCTTTGACGT TTTGAACTAA TCCCACTACC AAGTGCA | | | | |
| 2591 | TGGAGTCCAC GTTCTTTAAT AGTGGACTCT TGTTCCAAAC TGGAAACA CA CTCAACCCCTA TCCTCGGTCTA ACCTCAGGTG CAAGAAATTA TCACCTGAGA ACAAGGTTG ACCTTGTTGT GAGTTGGGAT AGAGCCAGAT | | | | |
| 2661 | TTCTTTTGAT TTATAAGGGA TTTTGCCGAT TTCGGCCTAT TGGTTAAAA ATGAGCTGAT TTAACAAAAA AAGAAAACTA AATATTCCTT AAAACGGCTA AAGCCGGATA ACCAATTTTT TACTCGACTA AATTGTTTTT | | | | |

113 / 154

```
2731 TTTAACGCGA ATTTTAACAA AATATTAACG CTTACAATTT AGGTGGCACT TTTCGGGGAA ATGTGCGCGG
AAATTGCGCT TAAAAATTGTT TTATAATTGC GAATGTTAAA TCCACCCGTA AAAGCCCCCTT TACACGCGCC

2801 AACCCCTATT TGTTTATTTT TCTAAATACA TATCCGCTCA TGAGACAATA ACCCTGATAA
TTGGGGATAA ACAAAATAAA AGATTATGT AAGTTTATAC ATAGCGGAGT ACTCTGTTAT TGGGACTATT

2871 ATGCTTCAAT AATATTGAAA AAGGAAGAGT ATGAGTATTC AACATTTCCTG TGTCGCCCTTT ATTCCCTTTT
TACGAAGTTA TTATAACTTT TTCCTTCTCA TACTCATAAG TTGTAAAGC ACAGCGGAA TAAGGGAATA

2941 TTGCGGCATT TTGCCCTTCCT GTTTTGTCTC ACCCAGAAAC GCTGGTGAAG GTAAAAAGATG CTGAAGATCA
AACGCCGTAA AACGGAAGGA CAAAAACGAG TGGGTCTTTG CGACCACCTT CATTTCTAC GACTTCTAGT

3011 GTTGGGTGCA CGAGTGGGT ACATCGAACT GGATCTCAAC AGCGGTGAAG TCCTTGAGAG TTTTCGCCCC
CAACCCACGT GCTCACCCAA TGTAGCTTGA CCTAGAGTTG TCGCCATTCT AGGAACCTCTC AAAAGCGGGG

3081 GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCG TATATATATCC CGTATTGACG
CTTCTTGCAA AAGTTACTA CTCGTGAAA CTCGAGAGC TTTCAAGACG ATACACCGCG CCATAATAGG GCATAACTGC

3151 CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG GTTGAGTACT CACCAGTCAC
GGCCCGTTCT CGTTGAGCCA GCGGCGTATG TGATAAGAGT CTTACTGAAC CAACTCATGA GTGGTCAGTG

3221 AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTA TGCAGTGCTG CCATAACCAT GAGTGATAAC
TCTTTTCGTA GAATGCCCTAC CGTACTGTCA TTCTCTTAAT ACGTCACGAC GGTATTGGTA CTCACTATTG

3291 ACTGCGGCCA ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTG CACAACATGG
TGACGCCGGT TGAATGAAGA CTGTTGCTAG CCTCCTGGCT TCCTCGATTG GCGAAAAAAC GTGTTGTACC

3361 GGGATCATGT AACTCGCCTT GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA
CCCTAGTACA TTGAGCGGAA CTAGCAACCC TTGGCCTCGA CTTACTTCGG TATGGTTTGC TGCTCGCACT

3431 CACCACGATG CCTGTAGCAA TGGCAACAAC GTTGGCGAAA CTATTAACTG GCGAACTACT TACTCTAGCT
GTGGTGCTAC GGACATCGTT ACCGTTGTTG CAACGCCGTT GATAATTGAC CGCTTGATGA ATGAGATCGA

3501 TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC ACTTCTGCGC TCGGCCCTTC
AGGCCCGTTG TTAATTATCT GACCTACCTC GCCTATTTC ACGTCCCTGG TGAAGACGCG AGCCGGGAAG
```

FIG._44F

114 / 154

```

3571 CGGCTGGCTG GTTTATTGCT GATAAATCTG GAGCCGGTGA GCGTGGGTCT CGCGGTATCA TTGCAGCACT
      GCCGACCGAC CAAATAACGA CTATTTAGAC CTCGGCCACT CGCACCCAGA GCGCCATAGT AACGTCGTGA

3641 GGGGCCAGAT GGTAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGGA GTCAGGCAAC TATGGATGAA
      CCCCGGTCTA CCATTCCGGA GGGCATAGCA TCAATAGATG TGCTGCCCTT CAGTCCGTTG ATACCTACTT

3711 CGAAATAGAC AGATCGCTGA GATAGGTGCC TCACTGATTA AGCATTTGGTA ACTGTCAGAC CAAGTTTACT
      GCTTTATCTG TCTAGCGACT CTATCCACGG AGTGACTAAT TCGTAACCAT TGACAGTCTG GTTCAAAATGA

3781 CATATATACT TTAGATTGAT TTAAAACTTC ATTTTAAATT TAAAGGATC TAGGTGAAGA TCCTTTTGA
      GTATATATGA AATCTAACTA AATTTTGAAG TAAAAATTAA ATTTTCTTAG ATCCACTTCT AGGAAAAACT

3851 TAATCTCATG ACCAAAATCC CTTAACGTGA GTTTTCGTTT CACTGAGCGT CAGACCCCGT AGAAAAGATC
      ATTAGAGTAC TGGTTTTAGG GAATTGCAT CAAAAGCAAG GTGACTCGCA GTCTGGGGCA TCTTTTCTAG

3921 AAAGGATCTT CTTGAGATCC TTTTTCCTG CGCGTAATCT GCTGCTTGCA AACAAAAAAA CCACCCGTAC
      TTTCCTAGAA GAACCTCTAG AAAAAAGAC GCGCATAGA CGACGAACGT TTGTTTT TTTT GGTGGCGATG

3991 CAGCGGTGGT TTGTTTGCCG GATCAAGAGC TACCAACTCT TTTTCCGAAG GTAACTGGCT TCAGCAGAGC
      GTCGCCACCA AACAAACGGC CTAGTTCTCG ATGGTTGAGA AAAAGGCTTC CATTGACCGA AGTCGTCTCG

4061 GCAGATACCA AATACTGTCC TTCTAGTGTA GCCGTAGTTA GGCCACCACCT TCAAGAATCT TGTAGCACCG
      CGTCTATGGT TTATGACAGG AAGATCACAT CGGCATCAAT CCGGTGGTGA AGTCTTTGAG ACATCGTGGC

4131 CCTACATACC TCGCTCTGCT AATCCTGTTA CCAGTGGCTG CTGCCAGTGG CGATAAGTCG TGTCTTACC
      GGATGTATGG AGCGAGACGA TTAGGACAAAT GGTCAACCGAC GACGGTCACC GCTATTTCAGC ACAGAATGGC

4201 GGTTGGACTC AAGACGATAG TTACCGGATA AGCGCGAGCG GTCGGGCTGA ACGGGGGTTT CGTGACACACA
      CCAACCTGAG TTCTGCTATC AATGGCCTAT TCCGCGTCCG CAGCCCGACT TGCCCCCAA GCACGTGTGT

4271 GCCCAGCTTG GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG AGCTATGAGA AAGCGCCACG
      CGGGTCGAAC CTCGCTTGCT GGATGTGGCT TGACTCTATG GATGTCGCAC TCGATACTCT TTCGCGGTGC

4341 CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG GCAGGGTCGG AACAGGAGAG CGCACGAGGG
      GAAGGGCTTC CCTCTTTCCG CCTGTCCATA GGCCATTCCG CGTCCCAGCC TTGTCCTCTC GCGTGTCTCC
  
```

FIG. 44G

115 / 154

4411 AGCTTCCAGG GGAACGCGG TGGTATCTTT ATAGTCCTGT CGGGTTTCG CACCTCTGAC TTGAGCGTCTG
TCGAAGGTCC CCTTTGCGG ACCATAGAAA TATCAGGACA GCCCCAAGCG GTGGAGACTG AACTCGCAGC
4481 ATTTTGTGA TGCTCGTCAG GGGGGCGGAG CCTATGAAA AACGCCAGCA ACGGGCCCTT TTTACGGTTC
TAAAAACACT ACGAGCAGTC CCCCCGCTC GGATACCTTT TTGGGTCTGT TGGCCCGGAA AATGCCAAG
4551 CTGGCCCTTT GCTGGCCTTT TGCTCACATG TTCTTTCCCTG CGTTATCCCC TGATTCTGTG GATAACCGTA
GACCGGAAA CGACCGGAAA ACGAGTGATC AAGAAAGGAC GCAATAGGGG ACTAAGACAC CTATTGGCAT
4621 TTACCGCCTT TGAGTGAGCT GATACCGCTC GCCGACGCCG AACGACCGAG CGCAGCGAGT CAGTGAGCGA
AATGGCGGAA ACTCACTCGA CTATGGCGAG CGGCTCGGC TTGCTGGCTC GCGTCGCTCA GTCACCTCGCT
4691 GGAAGCGGAA GAGCGCCCAA TACGCAAAAC CCTCTCCCC GCGCGTTGGC CGATTCAATTA ATGCAGCTGG
CCTTCGCCCTT CTCGCGGGTT ATGCGTTTGG CGGAGAGGGG CGCGCAACCG GCTAAGTAAT TACGTCGACC
4761 CACGACAGGT TTCCCGACTG GAAAGCGGGC AGTGAGCGCA ACGCAATTAA TGTGAGTTAG CTCACCTCAT
GTGCTGTCCA AAGGCTGAC CTTTCGCCCG TCACTCGCGT TGCCTTAATT ACACTCAATC GAGTGAGTAA
4831 AGGCACCCCA GGCTTTACAC TTTATGCTTC CGGCTCGTAT GTTGTGTGGA ATTGTGAGCG GATAACAATT
TCCGTGGGGT CCGAAATGTG AAATACGAAG GCCGAGCATA CAACACACCT TAACACTCGC CTATTGTTAA
4901 TCACACAGGA AACAGCTATG ACCATGATTA GCCAAGCGC GCAATTAAAC CTCACTAAAG GGAACAAAAG
AGTGTGTCCT TTGTCGATAC TGGTACTAAT GCGGTTCCG CGTTAATTGG GAGTGATTTT CCTTGTGTTTC

BssHII

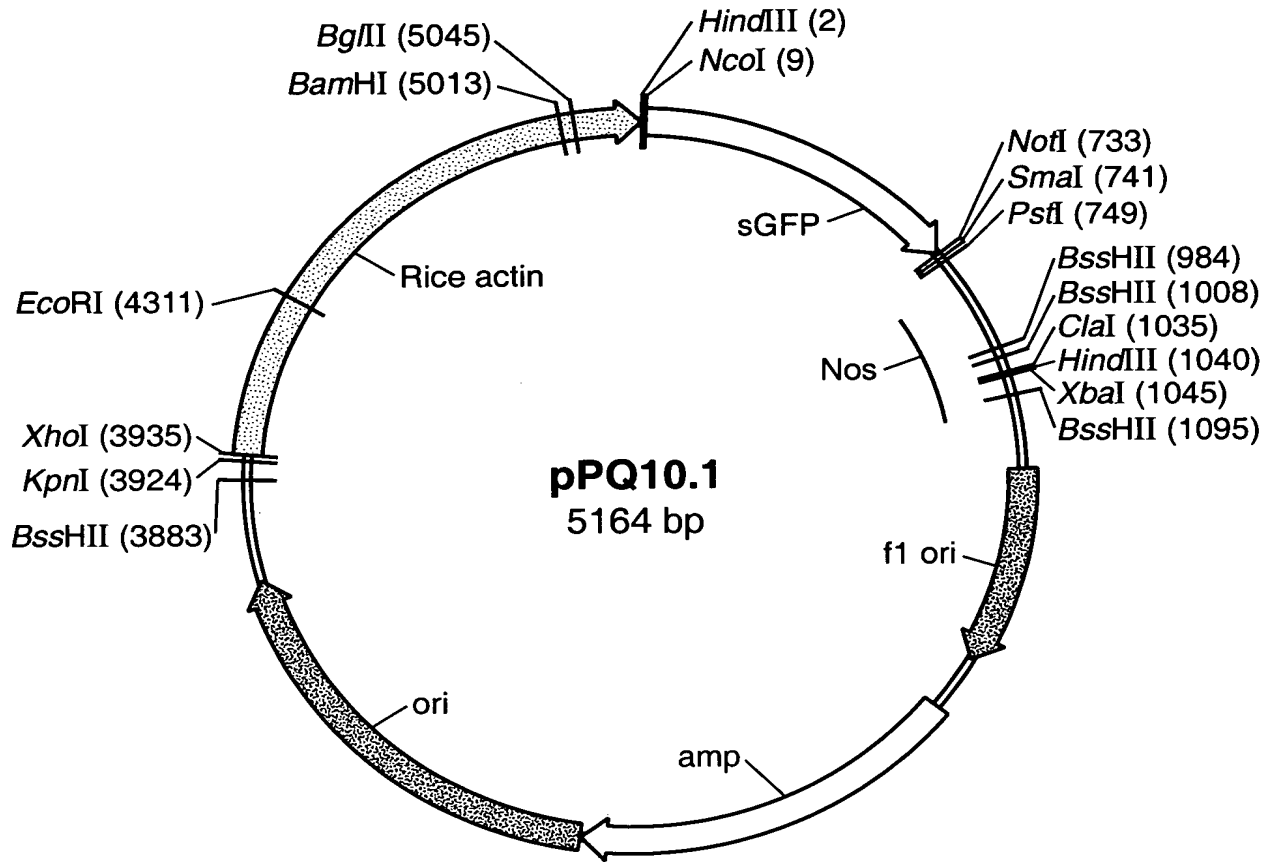
~~~~~

EcoR

4971 CTGG  
GACC

FIG.\_44H

116 / 154



**FIG. 45A**

117 / 154

```

HindIII NcoI
~~~~~
1 AAGCTTACCA TGGTGAGCAA GGGCGAGGAG CTGTTACCG GGTGGTGCC CATCTGGTC GAGCTGGACG
 TTCGAATGGT ACCACTCGTT CCCGCTCCTC GACAAGTGGC CCCACCACGG GTAGGACCAG CTCGACCTGC

71 GCGACGTGAA CGGCCACAAG TTCAGCGTGT CCGGCGAGGG CGAGGGCGAT GCCACCTACG GCAAGCTGAC
 CGCTGCACCTT GCCGGTGTTC AAGTCGCACA GCGCGTCCC GCTCCCGCTA CGGTGGATGC CGTTCGACTG

141 CCTGAAGTTC ATCTGCACCA CCGGCAAGCT GCCCGTGCCC TGGCCCCACCC TCGTGACCAC CTTCACCTAC
 GGACTTCAAG TAGACGTGGT GCGCGTTCGA CCGGCACGGG ACCGGGTGGG AGCACTGGTG GAAGTGGATG

211 GCGGTGCAGT GCTTCAGCCG CTACCCCGAC CACATGAAGC AGCAGCACTT CTTCAAGTCC GCCATGCCCCG
 CCGCACGTCA CGAAGTCGGC GATGGGCTG GTGTACTTCG TCGTGTGAA GAAGTTCAGG CCGTACGGGC

281 AAGGCTACGT CCAGGAGCGC ACCATCTTCT TCAAGGACGA CGGCAACTAC AAGACCCGCG CCGAGGTGAA
 TTCCGATGCA GGTCCCTCGC TGGTAGAAGA AGTTCCTGCT GCCGTTGATG TTCTGGGCGC GGCTCCACTT

351 GTTCGAGGCG GACACCCCTGG TGAACCGCAT CGAGCTGAAG GGCATCGACT TCAAGGAGGA CGGCAACATC
 CAAGCTCCCG CTGTGGGACC ACTTGGCGTA GCTCGACTTC CCGTAGCTGA AGTTCCTCCT GCCGTTGTAG

421 CTGGGGCACA AGCTGGAGTA CAACTACAAC AGCCACAACG TCTATATCAT GGCCGACAAG CAGAAGAACG
 GACCCCGTGT TCGACCTCAT GTTGATGTTG TCGGTGTTGC AGATATAGTA CCGCTGTTC GTCTTCTTGC

491 GCATCAAGGT GAATTCAAG ATCCGCCACA ACATCGAGGA CCGCAGCGTG CAGCTCGCCG ACCACTACCA
 CGTAGTTCCA CTTGAAGTTC TAGGCGGTGT TGTAGCTCCT GCCGTCGCAC GTCGAGCGGC TGGTGATGGT

561 GCAGAACACC CCCATCGGCG ACGGCCCCGT GCTGCTGCCC GACAACCACT ACCTGAGCAC CCAGTCCGCC
 CGTCTTGTGG GGTAGCCGC TGCCGGGGCA CGACGACGG CTGTTGGTGA TGGACTCGTG GGTGAGCGCG

631 CTGAGCAAAG ACCCCAACGA GAAGCGCGAT CACATGGTCC TGCTGGAGTT CGTGACCGCC GCCGGGATCA
 GACTCGTTTC TGGGGTTGCT CTTCCGCGTA GTGTACCAGG ACGACCTCAA GCACTGGCGG CGGCCCTAGT

```

FIG.\_45B

118 / 154

```

SmaI
~~~~~
NotI      PstI
~~~~~
701 CTCACGGCAT GGACGAGCTG TACAAGTAAA GCGCCCGCCC GGGCTGCAGG GAAACCACTG AAGGATGAGC
 GAGTGCCGTA CCTGCTCGAC ATGTTCAATT CGCCGGCGGG CCCGACGTCC CTTTGGTGAC TTCCCTACTCG

771 TGTAAGAAG CAGATCGTTC AAACATTGG CAATAAAGTT TCTTAAGATT GAATCCTGTT GCCGGTCTTG
 ACATTTCTTC GTCTAGCAAG TTTGTAAACC GTTATTTCAA AGAATTCTAA CTTAGGACAA CGGCCAGAAC

841 CGATGATTAT CATATAAATT CTGTTGAATT ACGTTAAGCA TGTAATAAAT AACATGTAAT GCATGACGTT
 GCTACTAATA GTATATTAAA GACAACTTAA TGCAATTCGT ACATTATTAA TTGTACATTA CGTACTGCAA

911 ATTTATGAGA TGGGTTTTTA TGATTAGAGT CCCGCAATTA TACATTTAAT ACGCGATAGA AAACAAAATA
 TAAATACTCT ACCCAAAAT ACTAATCTCA GGGCGTTAAT ATGTAAATTA TCGCGTATCT TTTGTTTTAT

 XbaI
                                ~~~~~
                                ClaI HindIII
                                ~~~~~
981 TAGCGCGCAA ACTAGGATAA ATTATCGCG CGGGTGTCAT CTATGTTACT AGATCGATAA GCTTCTAGAG
 ATCGCGCGTT TGATCCTATT TAATAGCGG CGCCACAGTA GATACAATGA TCTAGCTATT CGAAGATCTC

 BssHII
                                ~~~~~
1051  CGGCCGGTGG  AGTCCAATT  CGCCCTATAG  TGAGTCGTAT  TACGCGCGCT  CACTGGCCGT  CGTTTTACAA
    GCCGGCCACC  TCGAGGTTAA  GCGGGATATC  ACTCAGCATA  ATGCGCGCGA  GTGACCGGCA  GCAAAATGTT

1121  CGTCGTGACT  GGAATAACCC  TGGCGTTACC  CAACTTAATC  GCCTTGCAGC  ACATCCCCCT  TTCGCCAGCT
    GCAGCACTGA  CCTTTTGGG  ACCGCAATGG  GTTGAATTAG  CGGAACGTCG  TGTAGGGGGA  AAGCGGTCGA

1191  GGCGTAATAG  CGAAGAGGCC  CGCACCGATC  GCCCTTCCCA  ACAGTTGCGC  AGCCTGAATG  GCGAATGGGA
    CCGCATATATC  GCTTCTCCGG  GCGTGGCTAG  CGGGAAGGGT  TGTCAACGCG  TCGGACTTAC  CGCTTACCCT
```

FIG.\_45C

119 / 154

```

1261 CGCGCCCTGT AGCGGCGCAT TAAGCGGGC GGGTGTGGTG GTTACGGCA GCGTGACCGC TACACTTGCC
    GCGCGGGACA TCGCCGCGTA ATTCGCGCG CCCACACCAC CAATGCGCGT CGCACTGGCG ATGTGAACGG

1331 AGCGCCCTAG CGCCCGCTCC TTTTCGCTTC TTCTCGCCAC GTTCGCCGGC TTTCCCCCGTC
    TCGCGGGATC GCGGGCGAGG AAAGCGAAG AAGAGCGGTG CAAGCGGCCG AAAGGGGCAG

1401 AAGCTCTAAA TCGGGGGCTC CCTTTAGGGT TCCGATTAG TGCTTTACGG CACCTCGACC CCAAAAAACT
    TTCGAGATTT AGCCCCCGAG GAAATCCCA AGGCTAAATC ACGAAATGCC GTGGAGCTGG GGTTTTGTGA

1471 TGATTAGGGT GATGGTTCAC GTAGTGGCC ATCGCCCTGA TAGACGGTTT TTCGCCCTTT GACGTTGGAG
    ACTAATCCCA CTACCAAGTG CATCACCCGG TAGCGGGACT ATCTGCCAAA AGCGGGGAAA CTGCAACCTC

1541 TCCACGTTCT TTAATAGTGG ACTCTTGTTC CAAACTGGAA CAACACTCAA CCCATCTCG GTCTATTCTT
    AGGTGCAAGA AATTATCACC TGAGAACAAG GTTTGACCTT GTTGTGAGTT GGGATAGAGC CAGATAAGAA

1611 TTGATTTATA AGGGATTTTG CCGATTTTCG CCTATTGGTT AAAAAATGAG CTGATTTAAC AAAAATTATA
    AACTAAATAT TCCCTAAAC GCGTAAAGCC GGATAACCAA TTTTTTACTC GACTAAATG TTTTAAATTT

1681 CGCGAATTTT AACAAAATAT TAACGCTTAC AATTAGGTG GCACTTTTCG GGGAAATGTG CGCGGAACCC
    GCGCTTAAAA TTGTTTTATA ATTGCGAATG TTAAATCCAC CGTGAAAAGC CCTTTTACAC GCGCCTTGGG

1751 CTATTTGTTT ATTTTCTTAA ATACATTCAA ATATGTATCC GCTCATGAGA CAATAACCCCT GATAAATGCT
    GATAAACAAA TAAAAAGATT TATGTAAGTT TATACATAGG CGAGTACTCT GTTATTGGGA CTATTTACGA

1821 TCAATAATAT TGAAAAAGGA AGAGTATGAG TATTCAACAT TTCCCGTGTG CCCTTATTCC CTTTTTTGCG
    AGTTATTATA ACTTTTTCCT TCTCATACTC ATAAGTTGTA AAGGCACAGC GGAATAAAGG GAAAAAACGC

1891 GCATTTTGCC TTCCCTGTTT TGCTCACCCA GAAACGCTGG TGAAAGTAA AGATGCTGAA GATCAGTTGG
    CGTAAACCGG AAGGACAAAA ACGAGTGGGT CTTTGCAGC ACTTTCATTT TCTACGACTT CTAGTCAACC

1961 GTGCACGAGT GGGTTACATC GAACTGGATC TCAACAGCGG TAAGATCCTT GAGAGTTTTT GCCCCGAAGA
    CACGTGCTCA CCCAATGTAG CTTGACCCTAG AGTTGTGCCC ATCTTAGGAA CTCTCAAAAG CGGGGCTTCT

2031 ACGTTTTCCT ATGATGAGCA CTTTTAAAGT TCTGCTATGT GCGCGGGTAT TATCCCGTAT TGACGCCGGG
    TGCAAAAGGT TACTACTCGT GAAAATTTC AAGAGATACA CCGCGCCATA ATAGGCATA ACTGCGGCC

```

FIG. 45D

120 / 154

2101 CAAGAGCAAC TCGGTCGCCG CATACTAT TCTCAGAATG ACTTGGTTGA GTACTACCA GTACACAGAAA  
GTTCTCGTTG AGCCAGCGC GTATGTGATA AGAGTCTTAC TGAACCAACT CATGAGTGGT CAGTGTCTTT

2171 AGCATCTTAC GGATGGCATG ACAGTAAGAG AATTATGCAG TGCTGCCATA ACCATGAGTG ATAACACTGC  
TCGTAGAATG CCTACCGTAC TGTCATTCTC TTAATACGTC ACGACGGTAT TGGTACTCAC TATTGTGACG

2241 GGCCAACTTA CTTCTGACAA CGATCGGAGG ACCGAAGGAG CTAACCGCTT TTTTGCACAA CATGGGGGAT  
CCGGTTGAAT GAAGACTGTT GCTAGCCTCC TGGCTTCCTC GATTGGCGAA AAAACGTGTT GTACCCCTTA

2311 CATGTAACTC GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC AAACGACGAG CGTGACACCA  
GTACATTGAG CGGAACCTAGC AACCTTGSC CTCGACTTAC TTCGGTATGG TTTGCTGCTC GCACTGTGGT

2381 CGATGCCCTGT AGCAATGGCA ACAACGTTGC GCAAACCTAT AACTGGCGAA CTACTTACTC TAGCTTCCC  
GCTACGGACA TCGTTACCGT TGTGCAACG CGTTGATAA TTGACCGCTT GATGAATGAG ATCGAAGGGC

2451 GCAACAAATTA ATAGACTGGA TGGAGGCGGA TAAAGTTGCA GGACCACTTC TGCGCTCGGC CCTTCCGGCT  
CGTTGTTAAT TATCTGACCT ACCTCCGCCCT ATTTCAACGT CCTGGTGAAG ACGCGAGCCG GGAAGGCCGA

2521 GGCTGGTTTA TTGCTGATAA ATCTGGAGCC GGTGAGCGTG GGTCTCGCG TATCATTGCA GCACCTGGGC  
CCGACCAAT AACGACTATT TAGACCTCGC CCACTCGCAC CCAGAGCGCC ATAGTAACGT CGTGACCCCG

2591 CAGATGGTAA GCCCTCCCGT ATCGTAGTTA TCTACACGAC GGGGAGTCAG GCAACTATGG ATGAACGAAA  
GTCCTACCAT CGGGAGGCA TAGCATCAAT AGATGTGCTG CCCCTCAGTC CGTTGATACC TACTTGCCTT

2661 TAGACAGATC GCTGAGATAG GTGCCCTCACT GATTAAGCAT TGGTAACTGT CAGACCAAGT TTAATCATAT  
ATCTGTCTAG CGACTCTATC CACGGAGTGA CTAATTCGTA ACCATTGACA GTCTGGTTCA AATGAGTATA

2731 ATACTTTAGA TTGATTTAAA ACTTCATTTT TAATTTAAAA GGATCTAGGT GAAGATCCTT TTTGATAATC  
TATGAAATCT AACTAAATTT TGAAGTAAA ATTAATTTT CCTAGATCCA CTTCTAGGAA AAATAATTAG

2801 TCATGACCAA AATCCCTTAA CGTGAGTTT CGTTCCACTG AGCGTCAGAC CCCGTAGAAA AGATCAAAGG  
AGTACTGGTT TTAGGGAATT GCACTCAAAA GCAAGGTGAC TCGCAGTCTG GGGCATCTTT TCTAGTTTCC

2871 ATCTTCTTGA GATCCTTTT TTCTGCGCGT AATCTGCTGC TTGCAACAA AAAAACCAAC GCTACCAGCG  
TAGAAGAACT CTAGGAAAAA AAGACGCGCA TTAGACGACG AACGTTTGTT TTTTGGTGG CGATGGTCCG

FIG.-45E



121 / 154

```

2941 GTGGTTTGTG TGCCGGATCA AGAGCTACCA ACTCTTTTTC CGAAGGTAACT TGGCTTCAGC AGAGCGCAGA
CACCAAAACAA ACGGCCTAGT TCTCGATGGT TGAGAAAAAG GCTTCCATTG ACCGAAAGTCG TCTCGCGTCT

3011 TACCAAAATAC TGTCTTCTTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG CACCGCCTAC
ATGGTTTATG ACAGGAAGAT CACATCGGCA TCAATCCGGT GGTGAAGTTC TTGAGACATC GTGGCGGATG

3081 ATACCTCGCT CTGCTAATCC TGTTACCAGT GGCTGCTGCC AGTGGCGATA AGTCGTGTCT TACCGGGTTG
TATGGAGCGA GACGATTAGG ACAATGGTCA CCGACGACGG TCACCGCTAT TCAGCACAGA ATGGCCCCAAC

3151 GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTCCG GCTGAACGGG GGGTTCGTGC ACACAGCCCCA
CTGAGTTCTG CTATCAATGG CCTATTCCGC GTCGCCAGCC CGACTTGCCC CCCAAGCACG TGTGTGCGGT

3221 GCTTGGAGCG AACGACCTAC ACCGAACCTGA GATACCTACA GCGTGAGCTA TGAGAAAGCG CCACGCTTCC
CGAACCTCGC TTGCTGGATG TGGCTTGACT CTATGGATGT CGCACTCGAT ACTCTTTCGC GGTGCGAAGG

3291 CGAAGGGAGA AAGCGCGACA GGTATCCGGT AAGCGGCAGG GTCGGAACAG GAGAGCGCAC GAGGGAGCTT
GCTTCCCTCT TTCCGCCCTGT CCATAGGCCA TTCGCCGTCC CAGCCTTGTC CTCTCGCGTG CTCCCTCGAA

3361 CCAGGGGGAA ACGCCTGGTA TCTTTATAGT CCTGTGCGGT TTCCGCCACCT CTGACTTGAG CGTCGATTTT
GGTCCCCCTT TCGCGACCAT AGAAATATCA GGACAGCCCCA AAGCGGTGGA GACTGAACTC GCAGCTAAAA

3431 TGTGATGCTC GTCAGGGGGG CGGAGCCTAT GGA AAAACGC CAGCAACGCG GCCTTTTTC GGTTCCTGGC
AACTACGAG CAGTCCCCC GCCTCGGATA CCTTTTTCG GTCGTTGCGC CGGAAAAATG CCAAGGACCG

3501 CTTTGTGCTG CTTTGTGCTC ACATGTTCTT TCCTGCGTTA TCCCTTGATT CTGTGGATAA CCGTATTACC
GAAAAACGAC GGAAACGAG TGTACAAGAA AGGACGCAAT AGGGACTAA GACACCTATT GGCATAATGG

3571 GCCTTTGAGT GAGCTGATAC CGCTCGCCGC AGCCGAACGA CCGAGCGCAG CGAGTCAGTG AGCGAGGAAG
CGGAAACTCA CTCGACTATG GCGAGCGGCG TCGGCTTGCT GGCTCGCGTC GCTCAGTCAC TCGCTCCCTTC

3641 CGGAAGAGCG CCAATACGC AAACCGCCTC TCCCGCGCGG TTGGCCGATT CATTAATGCA GCTGGCACGA
GCCTTCTCGC GGGTTATGCG TTTGGCGGAG AGGGCGCGC AACCGGCTAA GTAATTACGT CGACCGTGCT

3711 CAGGTTTCCC GACTGGAAG CGGGCAGTGA GCGCAACGCA ATTAATGTGA GTTAGCTCAC TCATTAGGCA
GTCCAAAGGG CTGACCTTTC GCCCGTCACT CGCGTTGCGT TAATTACACT CAATCGAGTG AGTAATCCGT
  
```

FIG.\_45F

122 / 154

```

3781  CCCACAGGCTT  TACACTTTAT  GCTTCCGGCT  CGTATGTTGT  GTGGAATTGT  GAGCGGATAA  CAATTTCACA
      GGGGTCCGAA  ATGTGAAATA  CGAAGGCCGA  GCATACAACA  CACCTTAACA  CTCGCCATT  GTTAAAGTGT

      BssHII
      ~~~~~
3851 CAGGAAACAG CTATGACCAT GATTACGCCA AGCGGCAAT TAACCCTCAC TAAAGGGAAC AAAAGCTGGG
 GTCCTTTGTC GATACTGGTA CTAATGCGGT TCGCGGTTA ATTGGGAGTG ATTTCCCTTG TTTTCGACCC

 KpnI
      ~~~~~
3921  TACCGGGCCC  CCCCTCGAGG  TCATTTCATAT  GCTTGAGAAG  AGAGTCGGGA  TAGTCCAAAA  TAAACAAAAG
      ATGGCCCGGG  GGGGAGCTCC  AGTAAGTATA  CGAACTCTTC  TCTCAGCCCT  ATCAGGTTTT  ATTTTGTTTC

3991  GTAAGATTAC  CTGGTCAAAA  GTGAAAACAT  CAGTTAAAAG  GTGGTATAAG  TAAAATATCG  GTAATAAAAAG
      CATTCTAATG  GACCAGTTTT  CACTTTTGTA  GTCAATTTTC  CACCATATTC  ATTTTATAGC  CATTATTTTC

4061  GTGGCCCCAA  GTGAAATTTA  CTCCTTTTCTA  CTATTATAAA  AATTGAGGAT  GTTTTGTCGG  TACTTTTGATA
      CACCGGGTTT  CACTTTAAAT  GAGAAAAGAT  GATAATATTT  TTAACCTCCTA  CAAAACAGCC  ATGAAACTAT

4131  CGTCATTTT  GTATGAATTG  GTTTTAAAGT  TTATTCGCGA  TTTGGAAATG  CATATCTGTA  TTTGAGTCGG
      GCAGTAAAAA  CATACTTAAC  CAAAAATTCA  AATAAGCGCT  AAACCTTTAC  GTATAGACAT  AAACTCAGCC

4201  TTTTAAAGTT  CGTTGCTTTT  GTAAATACAG  AGGATTTGT  ATAAGAAATA  TC'TTAAAAA  ACCCATATGC
      AAAAAATCAA  GCAACGAAAA  CATTATGTC  TCCCTAAACA  TATTC'TTAT  AGAAATTTT  TGGGTATACG

      EcoRI
      ~~~~~
4271 TAATTGACA TAATTTTGA GAAAAATATA TATTCAGGCG AATCCACAA TGAACAATAA TAAGATTAAA
 ATTAAACTGT ATTAAAAACT CTTTTTATAT ATAAGTCCGC TTAAGGTGTT ACTTGTATT ATCTAAATT

4341 ATAGCTTGCC CCCGTTGCAG CGATGGGTAT TTTTCTAGT AAAATAAAAG ATAAACTTAG ACTCAAAACA
 TATCGAACGG GGGCAACGTC GCTACCCATA AAAAGATCA TTTTATTTTC TATTTGAATC TGAGTTTGT

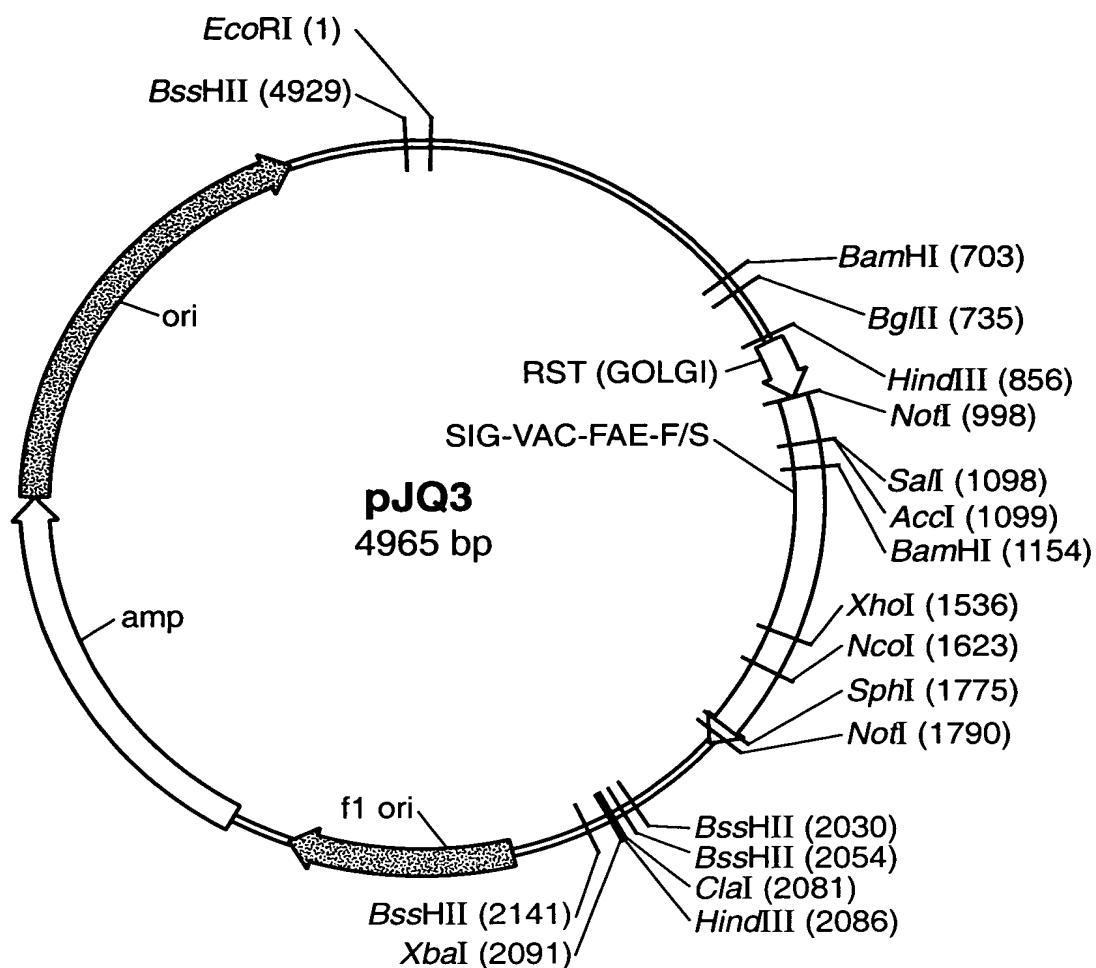
4411 TTTACAAAAA CAACCCCTAA AGTCCCTAAAG CCCAAAGTGC TATGCACGAT CCATAGCAAG CCCAGCCCCA
 AAATGTTTTT GTTGGGGATT TCAGGATTC GGGTTTCAG ATACGTGCTA GGTATCGTTC GGTCTGGGTT
```

FIG.\_45G

4481	CCCAACCCAA GGGTTGGGT	CCCAACCCAC GGGTGGGTG	CCAGTGCAG GGGTACCGT	CCAACTGGCA GGTTGACCGT	AAATGTCTCC TTATCAGAGG	ACCCCGGCA TGGGGCCCGT	CTATCACCGT GATAGTGGCA
4551	GAGTTGTCCG CTCAACAGGC	CACCACCGCA GTGGTGGCGT	CGTCTCGCAG GCAGAGCGTC	CCAAAAA GGTTTTTTTT	AAAAAGAAAG TTTTTCTTTC	AAAAAAAGA TTTTTTTTTCT	AAAAA TTTTCTTTTT
4621	CAGCAGGTGG GTCGTCCACC	GTCCGGGTG CAGGCCCAGC	TGGGGGCCGG ACCCCGGCC	AAAAGCAGG TTTTTCGTCC	AGGATCGCGA TCCTAGCGCT	GCAGCGACGA CGTCGCTGCT	GGCCCGGCC CCGGGCCGGG
4691	TCCCTCCGCT AGGAGGCGA	TCCAAAGAA AGGTTTCTTT	CGCCCCCCAT GCGGGGGTA	CGCCACTATA GCGGTGATAT	TACATACCCC ATGTATGGG	CCCCTCTCCT GGGAGAGGA	CCCATCCCC GGGTAGGGGG
4761	CAACCCCTACC GTTGGGATGG	ACCACACCA TGGTGGTGGT	CCACCACCTC GGTGGTGGAG	CTCCCCCCTC GAGGGGGAG	GCTGCCGGAC CGACGGCCTG	GACGAGCTCC CTGCTCGAGG	TCCCCCCTCC AGGGGGGAGG
4831	CCCTCCGCG GGGAGGCGC	CCGCCGGTAA GGCGGCCATT	CCACCCCGCC GGTGGGGCGG	CCTCTCCTCT GGAGAGGAGA	TTCTTTCTCTC AAGAAAGAGG	GTTTTTTTTT CAAAAAA	TCGTCTCGGT AGCAGAGCCA
4901	CTCGATCTTT GAGCTAGAA	GGCCTTGGTA CCGGAACCAT	GTTTGGGTGG CAAACCCACC	GCGAGAGCGG CGCTCTCGCC	CTTCGTCCGC GAAGCAGCGG	CAGATCGGTG GTCTAGCCAC	CGCGGGGAGG CGCCCTCCCC
4971	GCGGATCTC CGCCCTAGAG	GCGGCTGGCG CGCCGACCGC	TCTCCGGGCG AGAGGCCCGC	TGAGTCGGCC ACTCAGCCCG	CGGATCCTCG GCCTAGGAGC	CGGGGAATGG GCCCTTACC	GGCTCTCGGA CCGAGAGCCT
5041	TGTAGATCTT ACATCTAGAA	CTTTCTTTCT GAAAGAAAGA	TCTTTTGTG AGAAAAACAC	GTAGAAATTG CATCTTAAAC	AATCCCCTCAG TTAGGGAGTC	CATTGTTTCAT GTAACAAGTA	CGGTAGTTTT GCCATCAAAA
5111	TCTTTTCATG AGAAAAGTAC	ATTGTGACA TAAACACTGT	AATGCAGCCT TTACGTCGGA	CGTGC GGAGC GCACGCCCTC	TTTTTTGTAG AAAAAACATC	GTAG CATC	

**FIG. 45H**

124 / 154



**FIG. 46A**

125 / 154

```

ECORI
~~~~~
1  AATCCACAA TGAACAATAA TAAGATTAAA ATAGCTTGCC CCCGTTGCAG CGATGGGTAT TTTTCTTAGT
   TTAAGGTGTT ACTTGTTATT ATTCTAATTT TATCGAACGG GGGCAACGTC GCTACCCATA AAAAAGATCA

71  AAAATAAAAG ATAAACTTAG ACTCAAAACA TTTACAAAAA CAACCCCTAA AGTCTTAAAG CCCAAAGTGC
   TTTTATTTTC TATTGAATC TGAGTTTTGT AAATGTTTTT GTTGGGATT TCAGGATTTT CCGTTTCACG

141 TATGCACGAT CCATAGCAAG CCCAGCCCAA CCCAACCCAC CCAGTGCAG CCAACTGGCA
   ATACGTGCTA GGTATCGTTC GGGTCGGGTT GGGTTGGGT GGGTCACGTC GGTGACCGT

211 AATAGTCTCC ACCCCGGCA CTATCACCGT GAGTTGTCCG CACCACCGCA CGTCTCGAG CCAAAAAAAA
   TTATCAGAGG TGGGGCCGT GATAGTGCA CTCAACAGG GTGGTGGGT GCAGAGCGTC GGTTTTTTTT

281 AAAAAGAAAG AAAAAAAGA AAAAGAAAAA CAGCAGGTGG GTCCGGGTGG TGGGGGCCGG AAAAGCGAGG
   TTTTCTTTC TTTTCTTCT TTTTCTTCT TTTTCTTCT GTCTCCACC CAGGCCCAGC ACCCCCGGCC TTTTCGCTCC

351 AGGATCGCGA GCAGCGACGA GGGCCGGCCC TCCCTCCGCT TCCTAAAGAA CGCCCCCAT CGCCACTATA
   TCCTAGCGCT CGTCGTGCT CCGGGCCGGG AGGAGGCGA AGTTTCTTT GCGGGGGTA CCGGTGATAT

421 TACATACCCC CCCCTCTCCT CCCATCCCCC CAACCTACC ACCACACCA CCACCACTC CTCCCCCTC
   ATGTATGGGG GGGAGAGGA GGTAGGGGG GTTGGGATGG TGGTGGTGG GGTGGGGGAG GAGGGGGGAG

491 GCTGCCGGAC GACGAGCTCC TCCCCCTCC CCCTCCGCG CCGCCGGTAA CCACCCCGCC CCTCTCCTCT
   CGACGGCCTG CTGCTCGAG AGGGGGGAG GGGAGCGGC GCGGCCATT GGTGGGGCGG GGAGAGGAGA

561 TTCTTTCTCC GTTTTTTTT TCGTCTCGGT CTCGATCTTT GGCTTTGGTA GTTTGGGTGG GCGAGAGCGG
   AAGAAAGAG CAAAAAAA AGCAGAGCCA GAGCTAGAAA CCGGAACCAT CAAACCCACC CGCTCTCGCC

631 CTTCGTCCGC CAGATCGGTG CGCGGGAGGG GCGGATCTC GCGCTGGCG TCTCCGGGCG TGAGTCGGCC
   GAAGCAGCG GTCTAGCCAC GCGCCCTCCC CGCCCTAGAG CGCCGACCGC AGAGCCCGC ACTCAGCCCG

      BamHI
      ~~~~~
801 CGGATCCTCG CGGGGAATGG GGCTCTCGGA TGATAGTCTT CTTTCTTTCT TCTTTTGTG GTAGAATTG
 GCCTAGGAGC GCCCCTTACC CCGAGAGCCT ACATCTAGAA GAAAGAAAGA AGAAAAACAC CATCTTAAAC

 BglII
      ~~~~~

```

FIG.\_46B

126 / 154

```
771 AATCCCTCAG CATTGTTTCAT CGGTAGTTTT TCTTTTCATG ATTTGTGACA AATGCAGCCT CGTGCGGAGC
TTAGGGAGTC GTAACAAGTA GCCATCAAAA AGAAAAGTAC TAAACACTGT TTACGTCGGA GCACGCCCTCG

HindIII
~~~~~
841 TTTTGTGTAG GTAGAAGCTT ACCATGATCC ACACCAACCT CAAAAAGAAG TTCTCCCTCT TCATCCTCGT
AAAAAACATC CATCTTCGAA TGGTACTAGG TGTGGTTGGA GTTTTCTTTC AAGAGGGAGA AGTAGGAGCA

911 CTTCCTCCTC TTCGCCGTGA TCTGCGTGTG GAAGAAGGGC TCCGACTACG AGGCCCTCAC CCTCCAAGCC
GAAGGAGGAG AAGCGGCACT AGACGCACAC CTTCTTCCCG AGGCTGATGC TCCGGGAGTG GGAGGTTCCG

NotI
~~~~~
981 AAGGAGTTCC AAATGGCGGC CGCCTCCACG CAGGGCATCT CCGAAGACCT CTACAGCCGT TTAGTCGAAA
TTCCCTCAAGG TTTACCGCCG GCGGAGGTGC GTCCCGTAGA GGCTTCTGGA GATGTCGGCA AATCAGCTTT

SalI
~~~~~
AccI
~~~~~
1051 TGGCCACTAT CTCCCAAGCT GCCTACGCCG ACCTGTGCAA CATTCGTCG ACTATATATCA AGGAGAGAGAA
ACCGGTGATA GAGGTTTCGA CGGATGCGGC TGGACACGTT GTAAGGCAGC TGATAATAGT TCCCTCTCTT

BamHI
~~~~~
1121 AATTACAAAT TCTCAAACTG ACATTAAACGG ATGGATCCTC CGCGACGACA GCAGCAAAGA AATAATCACC
TTAAATGTTA AGAGTTTGAC TGTAATTGCC TACCTAGGAG GCGCTGCTGT CGTCGTTTCT TTATTAGTGG

1191 GTCTTCCGTG GCACTGGTAG TGATACGAAT CTACAACCTG ATACTAACTA CACCCTCACG CCTTTCGACA
CAGAAGGCAC CGTGACCATC ACTATGCTTA GATGTTGAGC TATGATTGAT GTGGGAGTGC GGAAGAGCTGT

1261 CCCTACCACA ATGCAACGGT TGTGAAGTAC ACGGTGGATA TTATATTGGA TGGGTCTCCG TCCAGGAGCA
GGGATGGTGT TACGTTGCCA ACACCTTCATG TGCCACCTAT AATATAACCT ACCCAGAGGC AGGTCTCTGT
```

FIG.- 46C

127 / 154

```
1331 AGTCGAGTCG CTTGTCAAAC AGCAGGTTAG CCAGTATCCG GACTACGGCG TGACCGTGAC CGGCCACKCC
 TCAGCTCAGC GAACAGTTTG TCGTCCAATC GGTCAATAGG CTGATGCGCG ACTGGCACTG GCCGGTGMGG

1401 CTCGGCGCCT CCCTGGCGGC ACTCACTGCC GCCAGCTGT CTGCGACATA CGACAAACATC CGCCTGTACA
 GAGCCGCGGA GGGACCGCGG TGAAGTACGG CGGGTCGACA GACGCTGTAT GCTGTTGTAG GCGGACATGT

 XhoI
                                ~~~~~
1471 CCTTCGGCGA ACCGCGCAGC GGCAATCAGG CCTTCGCGTC GTACATGAAC GATGCCCTTC AAGCCTCGAG
    GGAAGCCGCT TGGCGCGTCG CCGTTAGTCC GGAAGCGCAG CATGTACTTG CTACGGAAGG TTCGGAGCTC

1541 CCCAGATACG ACGCAGTATT TCCGGGTAC TCATGCCAAC GACGGCATCC CAAACCTGCC CCCGGTGGAG
    GGTCTATGC TCGTCAATAA AGCCCCAGTG AGTACGGTTG CTGCCGTAGG GTTTGGACGG GGGCCACCTC

                                NcoI
                                ~~~~~
1611 CAGGGGTACG CCCATGGCGG TGTAGAGTAC TGGAGCGTTG ATCCTTACAG CGCCCAGAAC ACATTTGTCT
 GTCCCCATGC GGTACCGCC ACATCTCATG ACCTCGCAAC TAGGAATGTC GCGGGTCTTG TGTAAACAGA

1681 GCACTGGGGA TGAAGTGCAG TGCTGTGAGG CCCAGGGCGG ACAGGGTGTG AATAATGCCG ACACGACTTA
 CGTGACCCCT ACTTCACGTC ACGACACTCC GGGTCCCGCC TGTCCCACAC TTATTACGCG TGTGCTGAAT

 SphI
                                ~~~~~
                                NotI
                                ~~~~~
1751 TTTTGGGATG ACGAGCGCG CATGCACCTG GCCGGTCGG GCGCGGAAA CCACTGAAGG ATGAGCTGTA
 AAAACCTAC TGCTCGCCG GTACGTGGAC CGCCACGCG CGCGCCTTT GGTGACTTCC TACTCGACAT

1821 AAGAAGCAGA TCGTTCAAAC ATTTGGCAAT AAAGTTTCTT AAGATTGAAT CCTGTGCGG GTCTTGCGAT
 TTCTTCGTCT AGCAAGTTTG TAAACCGTTA TTTCAAAGAA TTCTAACTTA GGACAAACGGC CAGAACGCTA

1891 GATTATCATA TAATTCTGT TGAATTACGT TAAGCATGTA ATAATTAACA TGTAAATGCAT GACGTTATTT
 CTAATAGTAT ATTAAAGACA ACTTAATGCA ATTCTGACAT TATTAATTGT ACATTACGTA CTGCAATAAA
```

FIG.-46D

128 / 154

```

1961 ATGAGATGGG TTTTATATGAT TAGAGTCCCG CAATTATACA TTTAATACGC GATAGAAAAC AAAATATAGC
 TACTCTACCC AAAAATACTA ATCTCAGGGC GTTAATATGT AAATTATGCG CTATCTTTTG TTTTATATCG
 BssHII
      ~~~
      XbaI
      ~~~~~

2031 GCGCAAACTA GGATAAATTA TCGCGCGCGG TGTCACTAT GTTACTAGAT CGATAAGCTT CTAGAGCGGC
 CGCGTTTGAT CCTATTTAAT AGCGCGCGCC ACAGTAGATA CAATGATCTA GCTATTGAA GATCTCGCCG
 BssHII
      ~~~
      ClaI HindIII
      ~~~~~

2101 CGGTGGAGCT CCAATTTCGCC CTATAGTGAG TCGTATTACG CGCGCTCACT GGCCGTCGTT TTACAACGTC
 GCCACCTCGA GGTAAAGCGG GATATCACTC AGCATAATGC GCGGAGTGA CCGGCAGCAA AATGTTGCAG

2171 GTGACTGGGA AAACCCCTGGC GTTACCCAAC TTAATCGCCT TGCAGCACAT CCCCCTTTTCG CCAGCTGGCG
 CACTGACCCCT TTTGGGACCG CAATGGGTTG AATTAGCGGA ACGTCGTGA GGGGAAAAGC GGTCGACCGC
 BssHII
      ~~~~~

2241  TAATAGCGAA GAGGCCCGCA CCGATCGCCC TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGGACGCG
      ATTATCGCTT CTCCGGGGCGT GGCTAGCGGG AAGGTTGTC AACGCGTCGG ACTTACCGCT TACCCCTGCGC

2311  CCCTGTAGCG GCGCATTAAG CGCGGCGGGT GTGGTGTTA CGCGCAGCGT GACCGCTACA CTTGCCAGCG
      GGGACATCGC CGCGTAATTC GCGCCGCCCA CACCACCAAT GCGCGTCGCA CTGGCGATGT GAACGGTCGC

2381  CCCTAGCGCC CGCTCCTTTC GCTTCTTCC CTTCCTTTCT CGCCACGTTT GCCGGCTTTC CCCGTCAAGC
      GGGATCGCGG GCGAGGAAAG CGAAAGAAGG GAAGGAAAGA GCGGTGCAAG CGGCCGAAAG GGGCAGTTTC

2451  TCCTAAATCGG GGGCTCCCTT TAGGGTTCCG ATTTAGTGCT TTACGGCACC TCGACCCCAA AAAACTTGAT
      AGATTTAGCC CCCGAGGGAA ATCCCAAGGC TAAATCACGA AATGCCGTGG AGCTGGGGTT TTTTGAACATA

2521  TAGGGTGATG GTTCACGTAG TGGGCCATCG CCCTGATAGA CGGTTTTTCG CCCTTTGACG TTGGAGTCCA
      ATCCCACTAC CAAGTGCATC ACCCGGTAGC GGGACTATCT GCCAAAAAGC GGGAAACTGC AACCTCAGGT
```

FIG.\_46E



129 / 154

2591 CGTCTCTTTAA TAGTGGACTC TTGTTCCAAA CTGGAACAAC ACTCAACCCCT ATCTCGGTCT ATTCCTTTGA  
GCAAGAAATT ATCACCTGAG AACAAAGTTT GACCTTGTTG TGAGTTGGGA TAGAGCCAGA TAAGAAAACT

2661 TTTATAAGGG ATTTTGCCGA TTTTCGGCCTA TTGGTTAAAA AATGAGCTGA TTTAACAAAA ATTTAACGCG  
AAATATTCCC TAAAACGGCT AAAGCCGGAT AACCAATTTT TTAATCGACT AAATTGTTTT TAAATTGCGC

2731 AATTTTAAACA AAATATTAAC GCTTACAATT TAGGTGGCAC TTTTCGGGGA AATGTGCGCG GAACCCCTAT  
TTAAAAATTGT TTTATATAATG CGAATGTTAA ATCCACCGTG AAAAGCCCTT TTACACGCGC CTTGGGGATA

2801 TTGTTTATTT TTCTAAAATAC ATTCAAAATAT GTATCCGCTC ATGAGACAAAT AACCTGATA AATGCTTCAA  
AACAAATAAA AAGATTATG TAAGTTTATA CATAGGCGAG TACTCTGTTA TTGGGACTAT TTACGGAAGTT

2871 TAATATTGAA AAAGGAAGAG TATGAGTATT CAACATTTC GTGTCGCCCT TATTCCTTT TTTGCGGCAT  
ATTATAACTT TTTCCCTTCTC ATACTCATAA GTTGTAAGG CACAGCGGGA ATAAGGGAAA AAACGCCGTA

2941 TTTGCCCTTC TGTTTTGTCT CACCCAGAAA CGCTGGTGAA AGTAAAAGAT GCTGAAAGATC AGTTGGGTGC  
AAACGGAAGG ACAAACGA GTGGTCTTT GCGACCACTT TCAATTTCTA CGACTTCTAG TCAACCCACG

3011 ACGAGTGGGT TACATCGAAC TGGATCTCAA CAGCGGTAAG ATCCTTGAGA GTTTTCGCCC CGAAGAACGT  
TGCTCACCCA ATGTAGCTTG ACCTAGAGTT GTCGCCATTC TAGGAACTCT CAAAAGCGGG GCTTCTTGCA

3081 TTTCCAATGA TGAGCACTTT TAAAGTTCTG CTATGTGGCG CGGTATTATC CCGTATTGAC GCCGGGCAAG  
AAAGGTTACT ACTCGTGAAA ATTTCAAGAC GATACACCGC GCCATAATAG GGCATAACTG CGGCCCGTTC

3151 AGCAACTCGG TCGCCGCATA CACTATTCTC AGAATGACTT GGTTGAGTAC TCACCAGTCA CAGAAAAAGCA  
TCGTTGAGCC AGCGCGTAT GTGATAAGAG TCTTACTGAA CCAACTCATG AGTGGTCAGT GTCCTTTTCGT

3221 TCTTACGGAT GGCATGACAG TAAGAGAAAT ATGCAGTGCT GCCATAACCA TGAGTGATAA CACTGCGGCC  
AGAAATGCCA CCGTACTGTC ATTCTCTTAA TACGTCACGA CCGTATTGGT ACTCACTATT GTGACGCCCCG

3291 AACTTACTTC TGACAAACGAT CGGAGGACCG AAGGAGCTAA CCGCTTTTTT GCACAACATG GGGGATCATG  
TTGAATGAAG ACTGTTGCTA GCCTCCTGGC TTCCCTCGATT GGCGAAAAAA CGTGTGTGAC CCCCTAGTAC

3361 TAACTCGCCT TGATCGTTGG GAACCGGAGC TGAATGAAGC CATACCAAAC GACGAGCGTG ACACCACGAT  
ATTGAGCGGA ACTAGCAACC CTTGGCCCTCG ACTTACTTCG GTATGGTTTG CTGCTCGCAC TGTGGTGCTA

FIG.\_46F

130 / 154

3431 GCCTGTAGCA ATGGCAACAA CGTTGCGCAA ACTATTAACT GCGAACTAC TTACTCTAGC TTCCCCGGCAA  
 CGGACATCGT TACCGTTGTT GCAACGCGTT TGATAATTGA CCGCTTGATG AATGAGATCG AAGGGCCGTT

3501 CAATTAATAG ACTGGATGGA GCGGATATAA GTTGCAGGAC CACTTCTGCG CTCGGCCCCCTT CCGGCTGGCT  
 GTTAATTATC TGACCTACCT CCGCCTATTT CAACGTCCCTG GTGAAGACGC GAGCCGGGAA GCGCGACCGA

3571 GGTTTATTGC TGATAAATCT GGAGCCGGTG AGCGTGGGTC TCGCGGTATC ATTGCAGCAC TGGGGCCAGA  
 CCAATAACG ACTATTTAGA CCTCGGCCAC TCGCACCCAG AGCGCCATAG TAACGTCTGT ACCCCGGTCT

3641 TGGTAAGCCC TCCCGTATCG TAGTTATCTA CACGACGGGG AGTCAGGCAA CTATGGATGA ACGAAATAGA  
 ACCATTCCGG AGGCATAGC ATCAATAGAT GTGCTGCCCC TCAGTCCGTT GATACCTACT TGCTTTATCT

3711 CAGATCGCTG AGATAGGTGC CTCACCTGATT AAGCATTTGGT AACGTGTCAGA CCAAGTTTAC TCATATATAC  
 GTCTAGCGAC TCTATCCACG GAGTGACTAA TTTCGTAACCA TTGACAGTCT GGTTCAAAATG AGTATATATG

3781 TTTAGATTGA TTTAAAACCTT CATTTTAAAT TTTAAAAGGAT CTAGGTGAAG ATCCTTTTGG ATAATCTCAT  
 AAATCTAACT AAATTTTGA GTAAAAATTA AATTTCCTA GATCCACTTC TAGGAAAAAC TATTAGAGTA

3851 GACCAAAATC CCTTAACGTG AGTTTTCGTT CCACTGAGCG TCAGACCCCG TAGAAAAAGAT CAAAGGATCT  
 CTGGTTTTAG GGAATTGCAC TCAAAAGCAA GGTGACTCGC AGTCTGGGC ATCTTTTCTA GTTTCCTAGA

3921 TCTTGAGATC CTTTTTTTCT GCGCGTAATC TGCTGCTTGC AAACAAAAAA ACCACCGCTA CCAGCGGTGG  
 AGAACTCTAG GAAAAAAGA CGCGCATTAG ACGACGAACG TTTGTTTTTT TGGTGGCGAT GGTCCGCCAC

3991 TTTGTTTGCC GGATCAAGAG CTACCAACTC TTTTTCGAA GGTAACCTGGC TTCAGCAGAG CGCAGATACC  
 AAACAAACGG CCTAGTTCTC GATGGTTGAG AAAAAGGCTT CCATTGACCG AAGTCGTCTC GCGTCTATGG

4061 AAATACTGTC CTTCTAGTGT AGCCGTAGTT AGGCCACCAC TTCAAGAACT CTGTAGCACC GCCTACATAC  
 TTTATGACAG GAAGATCACA TCGGCATCAA TCCGGTGGTG AAGTTCTTGA GACATCGTGG CCGATGTATG

4131 CTCGCTCTGC TAATCCTGTT ACCAGTGGCT GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGGTTGGACT  
 GAGCGAGACG ATTAGGACAA TGGTCACCGA CGACGGTCAC CGCTATTTCAG CACAGAAATGG CCCAACCTGA

4201 CAAGACGATA GTTACC GGAT AAGGCGCAGC GGTGCGGCTG AACGGGGGGT TCGTGCACAC AGCCCAGCTT  
 GTTCTGCTAT CAATGGCCTA TTCCGCGTCTG CCAGCCCGAC TTGCCCCCCA AGCACGTGTG TCGGGTCCGA

FIG.\_46G

131 / 154

4271 GGAGCGAACG ACCTACACCG AACTGAGATA CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCCGAA  
CCTCGCTTGC TGGATGTGGC TTGACTCTAT GGATGTCGCA CTCGATACTC TTTCGCGGTG CGAAGGGCTT

4341 GGGAGAAAGG CGGACAGGTA TCCGGTAAGC GGCAGGGTTCG GAACAGGAGA GCGCACGAGG GAGCTTCCAG  
CCCTCTTTCC GCCTGTCCAT AGGCCATTTCG CCGTCCCAGC CTTGTCCCTCT CGCGTGTCTCC CTCGAAGGTC

4411 GGGGAAACGC CTGGTATCTT TATAGTCCTG TCGGGTTTCG CCACCTCTGA CTTGAGCGTC GATTTTGTG  
CCCCTTTTCG GACCATAGAA ATATCAGGAC AGCCCAAAGC GGTGGAGACT GAACCTCGCAG CTAAAAACAC

4481 ATGCTCGTCA GGGGGGCGGA GCCTATGGAA AAACGCCAGC AACGCGGCCT TTTTACGGTT CCTGGCCCTTT  
TACGAGCAGT CCCCCCGCCT CGGATACCTT TTTGCGGTTCG TTGCGCCGGA AAAATGCCAA GGACCGGAAA

4551 TGCTGGCCCTT TTGCTCACAT GTTCTTTCTT GCGTTATCCC CTGATTCTGT GGATAAACCGT ATTACCGCCT  
ACGACCGGAA AACGAGTGTA CAAGAAAGGA CGCAATAGGG GACTAAGACA CCTATTGGCA TAATGGCGGA

4621 TTGAGTGAGC TGATACCGCT CGCCGCAGCC GAACGACCGA GCGCAGCGAG TCAGTGAGCG AGGAAAGCGGA  
AACTCACTCG ACTATGGCGA GCGGCGTTCG CTTGCTGGCT CGCGTCGCTC AGTCACTCGC TCCTTCGCCT

4691 AGAGCGCCCA ATACGCCAAC CGCCTCTCCC CGCGCGTTGG CCGATTCAAT AATGCAGCTG GCACGACAGG  
TCTCGCGGGT TATGCGTTTG GCGGAGAGGG GCGCGCAACC GGCTAAGTAA TTACGTCGAC CGTGCTGTCC

4761 TTTCCCGACT GGAAAGCGGG CAGTGAGCGC AACGCAATTA ATGTGAGTTA GCTCACTCAT TAGGCACCCC  
AAAGGCTGA CCTTTCGCCC GTCACTCGCG TTGCGTTAAT TACACTCAAT CGAGTGAGTA ATCCGTGGGG

4831 AGGCTTTACA CTTTATGCTT CCGGCTCGTA TGTGTGTGG AATTGTGAGC GGATAACAAT TTCACACACAGG  
TCCGAAATGT GAAATACGAA GGCCGAGCAT ACAACACACC TTAACACTCG CCTATTGTTA AAGTGTGTCC

4901 AACAGCTAT GACCATGATT ACGCCAAGCG CGCAATTAAC CCTCACTAAA GGGAAACAAA GCTGG  
TTTGTCGATA CTGGTACTAA TCGGGTTCGC GCGTTAATTG GGAGTGATTT CCCTTGTTTT CGACC

ECORI

BssHII

~~~~~

FIG._46H

132 / 154

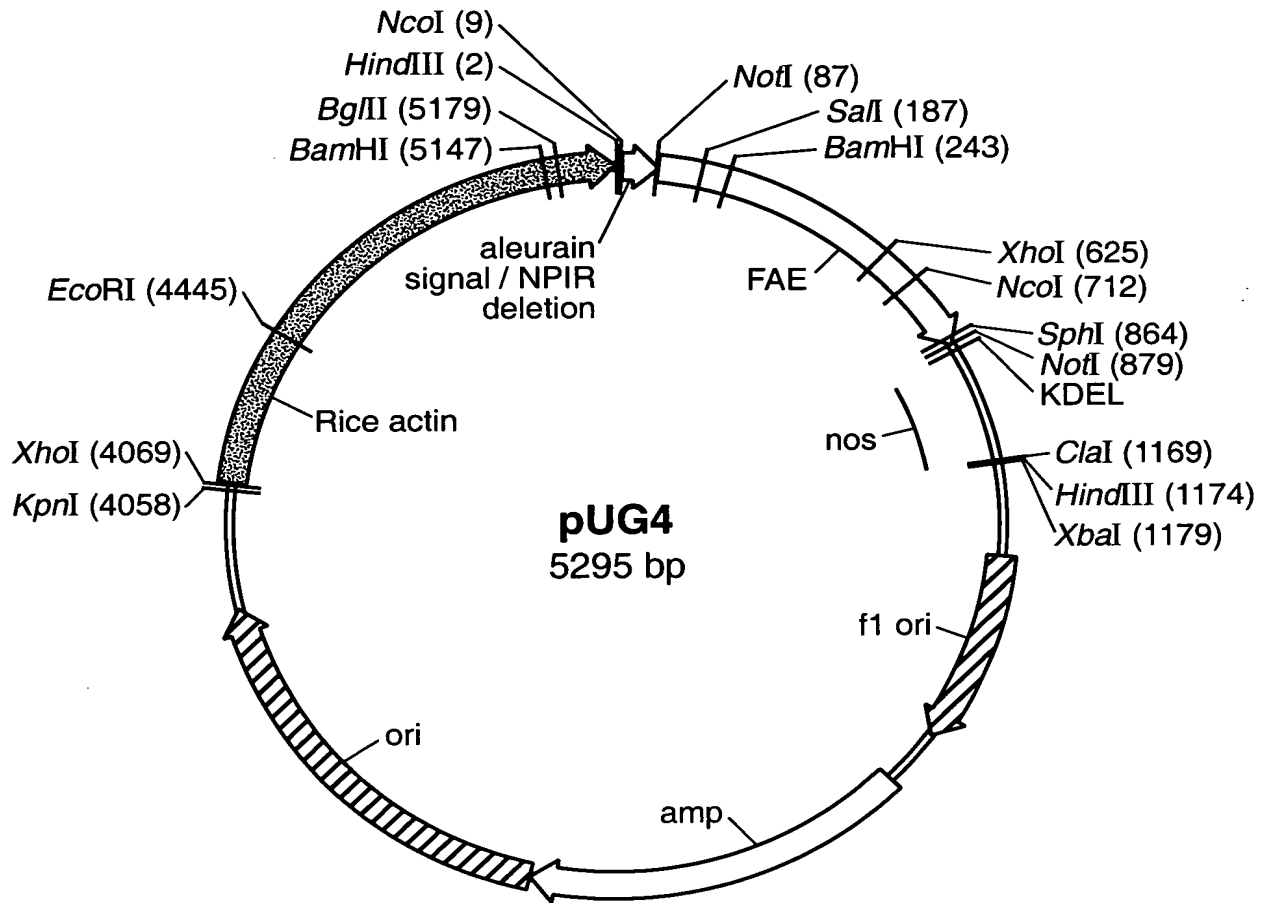


FIG. 47A

133 / 154

```

NGOI
~~~~~
HindIII
~~~~~
      M A H A R V L L L A L A V L A T A A V A V
1 AAGCTTACCA TGGCCACAGC CCGGTGCTC CTCCTGGGCG TCGCCGTGCT GCCACGGCC GCCGTGCGCG

      NotI
      ~~~~~
      . A S S R A A A S T Q G I S E D L Y S R L V E M .
71 TCGCCTCCTC CCGCGCGGCC GCCTCCACGC AGGCATCTC CGAAGACCTC TACAGCCGTT TAGTCGAAAT

      SalI
      ~~~~~
      . A T I S Q A A Y A D L C N I P S T I I K G E K
141 GGCCACTATC TCCCAAGCTG CCTACGCCGA CCTGTGCAAC ATTCGTCGA CTATTATCAA GGGAGAGAAA

      BamHI
      ~~~~~
      I Y N S Q T D I N G W I L R D D S S K E I I T V
211 ATTACAATT CTCAAACTGA CATTAACGGA TGGATCCTCC GCGACGACAG CAGCAAAGAA ATAATCACCG
      . F R G T G S D T N L Q L D T N Y T L T P F D T .
281 TCTTCCGTGG CACTGGTAGT GATACGAATC TACAATCGA TACTAACTAC ACCCTCACGC CTTTCGACAC
      . L P Q C N G C E V H G G Y Y I G W V S V Q D Q
351 CCTACCACAA TGCAACGGTT GTGAAGTACA CGGTGGATAT TATATTGGAT GGGTCTCCGT CCAGGACCAA
      V E S L V K Q Q V S Q Y P D Y A L T V T G H X L
421 GTCGAGTCGC TTGTCAAACA GCAGGTTAGC CAGTATCCGG ACTACGCGCT GACCGTGACC GGCCACKCCC
      . G A S L A A L T A A Q L S A T Y D N I R L Y T .
491 TCGGCGCCTC CCTGGCGGCA CTCACTGCCG CCCAGCTGTC TCGACATAC GACAACATCC GCCTGTACAC

      XhoI
      ~~~~~
      . F G E P R S G N Q A F A S Y M N D A F Q A S S
561 CTTGGCGGAA CCGCGCAGCG GCAATCAGGC CTTGCGTCTG TACATGAACG ATGCTTCCA AGCCTCGAGC
      P D T T Q Y F R V T H A N D G I P N L P P V E Q
631 CCAGATACGA CGCAGTATTT CCGGGTCACT CATGCCAACG ACGGCATCCC AAACCTGCCC CCGGTGGAGC

      NgoI
      ~~~~~
```

FIG..47B

134 / 154

```

      . G Y A H G G V E Y W S V D P Y S A Q N T F V C .
701 AGGGGTACGC CCATGGCGGT GTAGAGTACT GGAGCGTTGA TCCTTACAGC GCCAGAACA CATTTGTCTG
      . T G D E V Q C C E A Q G G Q G V N A H T T Y
771 CACTGGGGAT GAAGTGCAGT GCTGTGAGGC CCAGGGCGGA CAGGGTGTGA ATAATGCGCA CACGACTTAT

      SphI                               NotI
      ~~~~~                               ~~~~~
      F G M T S G A C T W P V A A A E P L K D E L *
841 TTTGGGATGA CGAGCGGCGC ATGCACCTGG CCGGTCGGG CCGCGGAACC ACTGAAGGAT GAGCTGTAAA
911 GAAGCAGATC GTTCAAAACAT TTGGCAATAA AGTTCTTAA GATTGAATCC TGTGCGCGGT CTTGCGGATGA
981 TTATCATATA ATTTCTGTTG AATTACGTTA AGCATGTAAT AATTAACATG TAATGCATGA CGTTATTATAT
1051 GAGATGGGTT TTTATGATTA GAGTCCCGCA ATTATACATT TAATACGCGA TAGAAAACAA AATATAGCGC

      HindIII
      ~~~~~

1121 GCAAAC TAGG ATAAATTATC GCGCGCGGTG TCATCTATGT TACTAGATCG ATAAGCTTCT AGAGCGGCGC
1191 GTGGAGCTCC AATTCGCCCT ATAGTGAGTC GTATTACGCG CGCTCACTGG CCGTCGTTTT ACAACGTCGT
1261 GACTGGGAAA ACCCTGGCGT TACCCAACTT AATCGCCTTG CAGCACATCC CCTTTCGCC AGCTGGCGTA
1331 ATAGCGAAGA GCGCCGCACC GATCGCCCTT CCCAACAGTT GCGCAGCCTG AATGGCGAAT GGGACGCGCC
1401 CTGTAGCGGC GCATTAAGCG CGGCGGGTGT GGTGGTTACG CGCAGCGTGA CCGCTACACT TGCCAGCGCC
1471 CTAGCGCCCG CTCCTTTTCG CTCTTTCCCT TTTCTTCCCT TTAGTGCTTT ACGGCACCTC GAGCTTCCCG CGTCAAGCTC
1541 TAAATCGGGG GCTCCCTTTA GGGTCCGAT TTAGTGCTTT GGTGACCTC GACCCCAAAA AACTTGATTA
1611 GGGTGATGGT TCACGTAGTG GGCCATCGCC CTGATAGACG GTTTTTTCGCC CTTTGACGTT GGAGTCCACG
1681 TTCTTTAATA GTGGACTCTT GTTCCAAACT GGAACAACAC TCAACCCCTAT CTCGGTCTAT TCTTTTGATT
1751 TATAAGGGAT TTTGCCGATT TCGGCCCTATT GGTAAATAA TGAGCTGATT TAACAAAAAT TTAACGCGAA
1821 TTTTAACAAA ATATTACGC TTACAATTTA GGTGGCACTT ATCCGCTCAT GAGACAATAA CCTGATAA TGCTTCAATA
1891 GTTTATTTTT CTAAATACAT TCAATATATG ATCCGCTCAT ACATTTCCGT GTCCGCCCTTA TTCCCTTTTT TGCGGCATTT
1961 ATATTGAAA AGGAAGAGTA TGAGTATTCA CCAGATAAAG CTGGTAAAG TAAAGATGAG TGAAGATCAG TTGGGTGCAC
2031 TGCCCTCCCTG TTTTGTGCTCA CCCAGAAAAC GATCTCAACA GCGGTAAGAT CCTTGAGAGT TTTCCGCCCG AAGAACGTTT
2101 GAGTGGGTTA CATCGAACTG AGCTCTTTTA AGTTCTGCT ATGTGGCGCG GTATTATCCC GTATTGACGC CGGGCAAGAG
2171 TCCAATGATG AGCACTTTTA CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTCACA GAAAAGCATC
2241 CAACTCGGTC GCCGATACA CATGACAGTA AGAGAATTAT GCAGTGCTGC CATAAACCATG AGTGATAACA CTGCGGCCAA
2311 TTACGGATGG CATGACAGTA AGAGAATTAT GCAGTGCTGC CATAAACCATG AGTGATAACA CTGCGGCCAA
2381 CTTACTTCTG ACAACGATCG GAGGACCGAA GGAGCTAACC GCTTTTTCG ACAACATGGG GGATCATGTA
2451 ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA TACCAAAACGA CGAGCGTGAC ACCACGATGC
2521 CTGTAGCAAT GGCAACAACG TTGCGCAAC TATTAACCTG CGAACTACTT ACTCTAGCTT CCCGGCAACA

```

FIG._47C

135 / 154

```

2591 ATTAATAGAC TGGATGGAG CGGATAAAGT TGCAGGACCA CTTCGCGCT CGCCCTTCC GCGTGGCTGG
2661 TTTATTGCTG ATAAATCTGG AGCCGGTGAG CGTGGGTCTC GCGGTATCAT TGCAGCACTG GGGCCAGATG
2731 GTAAGCCCTC CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC GAAATAGACA
2801 GATCGCTGAG ATAGGTGCTT TAAAACCTCA TTTTTCGTTCC ACTGAGCGTC AGACCCCGTA AAGGATCTTC
2871 TAGATTGATT TAAAACCTCA TTTTTCGTTCC ACTGAGCGTC AGACCCCGTA AAGGATCTTC
2941 CCAAAATCCC TTAACGTGAG TTTTTCGTTCC ACTGAGCGTC AGACCCCGTA AAGGATCTTC
3011 TTGAGATCCT TTTTTCGTTCC GCGTAATCTG GCGTAATCTG GCGTAATCTG GCGTAATCTG AAGGATCTTC
3081 TGTTTGCCCG ATCAAGAGCT ACCAATCTTT CCCTAGTTAG GCGTAATCTG GCGTAATCTG AAGGATCTTC
3151 ATACTGTCTT TCTAGTGTAG CCCTAGTTAG GCGTAATCTG GCGTAATCTG GCGTAATCTG AAGGATCTTC
3221 CGCTCTGCTA ATCTGTGTAG CCCTAGTTAG GCGTAATCTG GCGTAATCTG GCGTAATCTG AAGGATCTTC
3291 AGACGATAGT TACCGGATAA GCGTAATCTG GCGTAATCTG GCGTAATCTG GCGTAATCTG AAGGATCTTC
3361 AGCGAACGAC CTACACCGAA CTGAGATACC GCGTAATCTG GCGTAATCTG GCGTAATCTG AAGGATCTTC
3431 GAGAAAGGCG GACAGGTATC GCGTAATCTG GCGTAATCTG GCGTAATCTG GCGTAATCTG AAGGATCTTC
3501 GGAACGCTT GGTATCTTTA GGTATCTTTA GGTATCTTTA GGTATCTTTA GGTATCTTTA AAGGATCTTC
3571 GCTCGTCAGG GGGCGGAGC CTATGGAATA CTATGGAATA CTATGGAATA CTATGGAATA AAGGATCTTC
3641 CTGGCCCTTT GCTCACATGT GCTCACATGT GCTCACATGT GCTCACATGT GCTCACATGT AAGGATCTTC
3711 GAGTGAGCTG ATACCGCTCG ATACCGCTCG ATACCGCTCG ATACCGCTCG ATACCGCTCG AAGGATCTTC
3781 AGCGCCCAAT ACGCAACCG ACGCAACCG ACGCAACCG ACGCAACCG ACGCAACCG AAGGATCTTC
3851 TCCCGACTGG AAAGCGGCA GTGAGCGCA GTGAGCGCA GTGAGCGCA GTGAGCGCA AAGGATCTTC
3921 GCTTTACACT TTATGCTTCC TTATGCTTCC TTATGCTTCC TTATGCTTCC TTATGCTTCC AAGGATCTTC

3991 ACAGCTATGA CCATGATTAC GCCAAGCGCG CAATTAAACC TCACTAAAGG GAACAAAAGC TGGGTACCGG
      XhoI
      ~~~~~
4061 GCGCCCTCTC GAGGTCATTC ATATGCTTGA GAAGAGAGTC GGGATAGTCC AAAATAAAAC AAAGGTAAGA
4131 TTACCTGGTC AAAAGTGAAA ACATCAGTTA AAAGGTGGTA TAAGTAAAT ATCGGTAATA AAAGGTGGCC
4201 CAAAGTGAAA TTTACTCTTT TCTACTATTA TAAAATTTGA GGATGTTTGG TCGGTACTTT GATACGTCAT
4271 TTTTGTATGA ATTGGTTTTT AAGTTTATTC GCGATTGGGA AATGCATATC TGTATTGAG TCGGTTTTTA
4341 AGTTCTGTTG TTTTGTAAAT ACAGAGGGAT TTGTATAAGA AATATCTTTA AAAAAACCCAT ATGCTAATTT
      EcoRI
      ~~~~~
4411 GACATAAATT TTGAGAAAAA TATATATTCA GCGGAATTCC ACAATGAACA ATAATAAGAT TAAAAATAGCT
4481 TGCCCCCGTT GCAGCGATGG GTATTTTCTC TAGTAAAATA AAAGATAAAC TTAGACTCAA AACATTTACA
4551 AAAACAACCC CTAAAGTCCT AAAGCCCAAA GTGCTATGCA CGATCCATAG CAAGCCAGC CCAACCCCAAC
4621 CCAACCCCAAC CCACCCCACT GCAGCCCACT GGCAAAATAG CTCCACCCCT GGCATATCA CCGTGAGTTG

```

FIG._47D

```

4691 TCCGCACCAC CGCACGTCTC GCAGCCAAAA AAAAAGG AAGAAAAAGA AAAACAGCAG
4761 GTGGGTCCGG GTCGTGGGGG CCGGAAAAGC GAGGAGGATC GCGAGCAGCG ACGAGGCCCG GCCCTCCCTC
4831 CGCTTCCAAA GAAACGCCCC CCATCGCCAC TATATACATA CCCCCCCTC TCCTCCCATC CCCCCAACCC
4901 TACCACCACC ACCACCACCA CCTCCTCCCC CCTCGCTGCC GGACGACGAG CTCCTCCCCC CTCCTCCCTCC
4971 GCCGCCGCCG GTAACCAACC CGCCCCCTCTC CTCTTCTTTT CTCCGTTTTT TTTTTCGTCT CCGTCTCGAT
5041 CTTTGGCCCTT GGTAGTTTGG GTGGGCGAGA GCGCTTTCGT CCCCCAGATC GGTGCGCGGG AGGGCGGGGA
                                     BamHI
5111 TCTCGCGGCT GCGGTCCTCC GCGGTGAGTC GGCCCCGATC CTCGCGGGGA ATGGGGCTCT CGGATGTAGA
                                     ~~~~
5181 TCTTCTTTCT TCTTCTTTT TGTGGTAGAA TTTGAATCCC TCAGCATGTG TCATCGGTAG TTTTTCCTTT
5251 CATGATTGTG GACAAAATGA GCCTCGTGCG GAGCTTTTTT GTAGC

```

FIG. 47E

137 / 154

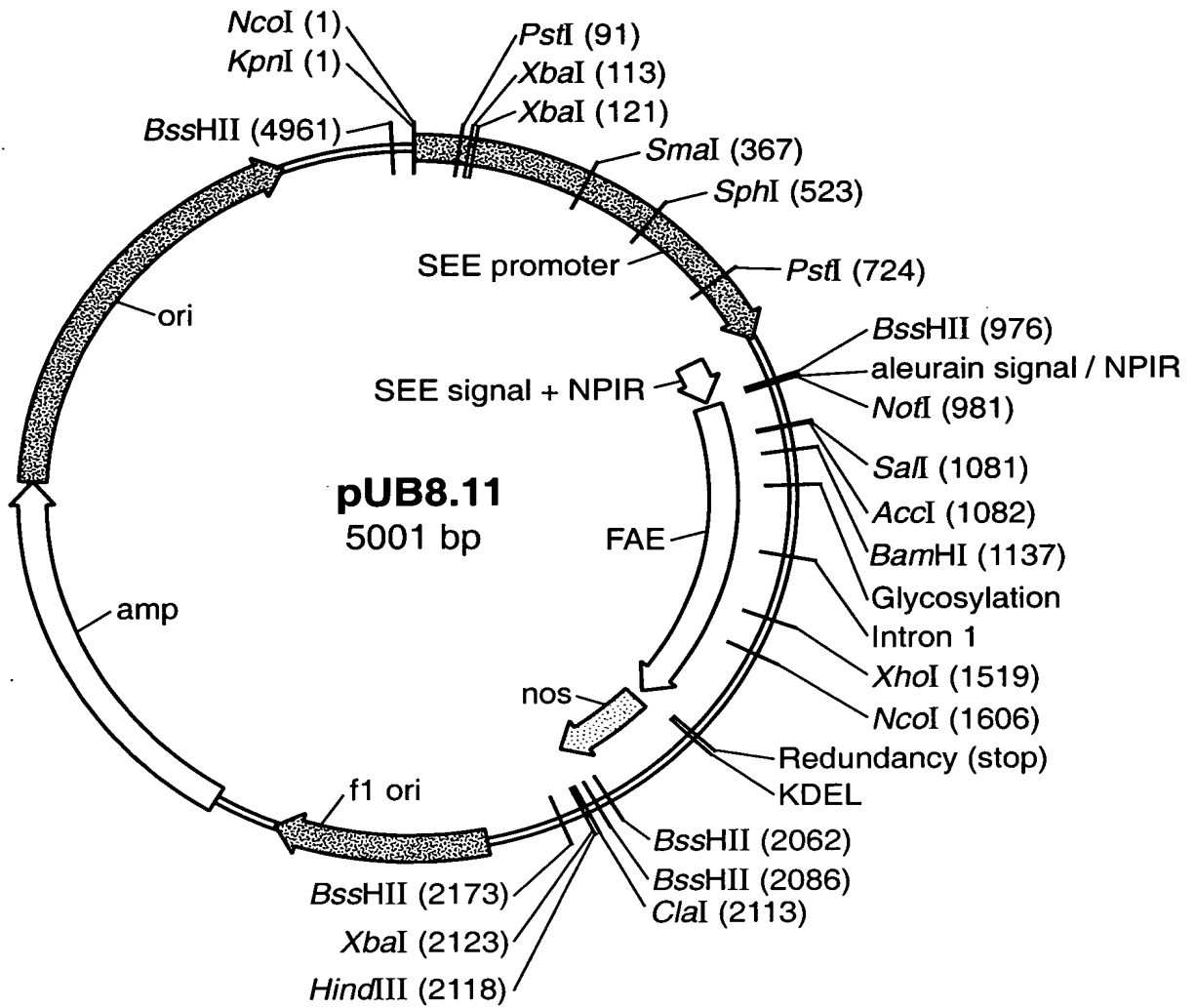


FIG._48A

138 / 154

```

NcoI
~~~~~
KpnI
~
1  CATGGGCCAG GTATAATTAT GGGATATCTC AAGCAAAATAA TCGAAATATC ACCATTGGCT ACAATATCTG
   GTACCCGGTC CATATTAATA CCTATAGAG TTCGTTTATT AGCTTTATAG TGGTAACCGA TGTATATAGAC

                                     PstI
                                     ~~~~~
71  AGCTCCGAGT TCTGACTGCA GTCTGGATGA CGCGTGTGTGT ATCTAGAACT CTAGATAGCA CAGCCACAGC
   TCGAGGCTCA AGACTGACGT CAGACCTACT GCGCACAAACA TAGATCTTGA GATCTATCGT GTCGGTGTCTG

                                     XbaI
                                     ~~~~~
141 ACCTACAGGA GTGCGACACT TGTGGACTGT AGTAGTGTG GAGACGGAGC TCTTTCCTTAC CTCTGACGCT
   TGGATGTCCT CACGCTGTGA ACACCTGACA TCATCACAAC CTCTGCCTCG AGAAAGGATG GAGGACTGCA

211 TGCCGCGCGTT GTCCATTCCA ACGGCATCAC TCCTCAACCAA TCACGCGCTC CCAACAACAAAT ATCGTCCCCC
   ACGGCGGCAA CAGGTAAGGT TGCCGTAGTG AGAGTTGGTT AGTGCGCGAG GGTGTGTTTA TAGCAGGGGG

281 ATGTCTTGGC GGAGAGAGAG TACATACATG CTGTGCGGCC GTTTTGTCT GAATCTCGCT TCCACTGGCC
   TACAGAACCG CCTCTCTCTC ATGTATGTAC GACAGCGCGG CAAAACACA CTTAGAGCGA AGGTGACCCG

                                     SmaI
                                     ~~~~~
351 AATCAGCTCA GCTCCCGGGA GCTCACTCAT TCAAGATCCC ATCGTCGTCTG TCACCCCTGG CGTCATGGGA
   TTAGTCGAGT CGAGGGCCCT CGAGTGAGTA AGTTCTAGGG TAGCAGCAGC AGTGGGGACC GCAGTACCCT

421 TGGAAAAGAA CCTCCGTTGC TCGGATGAGT CAGCCATATC CCCGAACAGA GTACTGCAAG ATAACCCAAT
   ACCTTTTCTT GGAGGCAACG AGCCTACTCA GTCGGTATAG GGGCTTGTCT CATGACGTTT TATTGGGTTA

                                     SphI
                                     ~~~~~
491 TCAGATTCCC CCAATAGAGA AAGTATAGCA TGCTTTCGGG TTTTGTGTTGG CTTAATTGAC TTTATTTTGG
   AGTCTAAGGG GGTATATCTCT TTCATATCGT ACGAAAGCCC AAAACAACCC GAATTAACTG AAATAAAAAAC

561 TTGGAGTTGA ATGCTGATTT GTTGTGTAAA ATGCCCAACC ATCTGAATAT CGAGACGGAT AATAGGCTGG
   AACCTCAACT TACGACTAAA CAACACATTT TACGGGTTGG TAGACTTATA GCTCTGCCTA TTATCCGACC
  
```

FIG._48B

139 / 154

```

631  CTAATTAATT TATAGCAAGA TTCTGTAGTG CACATCGCAA ATATCTTTCT GGCATTACA GCTGGAGGCT
      GATTAATTAA ATATCGTTCT AAGACATCAC GTGTAGCGTT TATAGAAAGA CCCGTAATGT CGACCTCCGA

      PstI
      ~~~~~
701  TCATCAGCCT GAAACACTCT GCAGAGCCTG AAGCAAAGTG TGAAGCGTGG CGATGAGATG GGTATAAAAC
      AGTAGTCGGA CTTTGTGAGA CGTCTCGGAC TTCGTTCAAC ACTTCGCACC GCTACTCTAC CCATATTTTG

771  CCCCCGCACC GGGACGCGAG CTCGCCCTTA CCAGTACCAT CTCGCCCTGC TCCCCCTGCC GGACGACCCA
      GGGGCCGTGG CCTGCGCTC GAGGCGGAT GGTCAATGTA GAGCGGAGCG AGGGGACGG CCTGCTGGGT

841  GTAAAATACT GTTGCCCACT CGCCGGCGAG ATGGCCACG GCCGCATCCT CTTCTTGGCG CTCGCCGTCT
      CATTTTATGA CAACGGGTGA GCGGCCGCTC TACCGGGTGC CGCGGTAGGA GAAGAACC GC GAGCGGCAGA

      BssHII
      ~~~~~
      NotI
      ~~~

911  TGGCCACCGC CGCGGTGGCC GCCGCATCNT TGGCGGACTC CAACCCGATC CGGCCCGTCA CCGAGCGCGC
      ACCGTTGGCG GCGCCACCGG CGCGGTAGNA ACCGCTGAG GTTGGCTAG GCCGGGCAGT GGCTCGCGCG

      NotI
      ~~~~~
981  GGCCGCCCTCC ACGCAGGGCA TCTCCGAAGA CCTCTACAGC CGTTTAGTGC AAATGGCCAC TATCTCCCAA
      CCGCGGAGG TGCCTCCCGT AGAGGCTTCT AGAGGCTGAG GCAATCAGC TTACCCGGTG ATAGAGGGTT

      Sali
      ~~~~~
      AccI
      ~~~~~

1051 GCTGCCCTACG CCGACCTGTG CAACATTCGG TCGACTATTA TCAAGGGAGA GAAAATTAC AATTCTCAA
      CGACGGATGC GGCTGGACAC GTTGTAAGGC AGCTGATAAT AGTTCCTCT CTTTTAAATG TTAAGAGTTT

```

FIG._48C

140 / 154

```

      BamHI
      ~~~~~
1121  CTGACATTAA CGGATGGATC CTCCGGGACG ACAGCAGCAA AGAAATAATC ACCGTCTTCC GTGGCAGTGG
      GACTGTAATT GCCTACCTAG GAGGCGCTGC TGTCTGTCGTT TCTTTATTAG TGGCAGAAGG CACCGTGACC

1191  TAGTGATACG AATCTACAAC TCGATACTAA CTACACCCCTC ACGCCTTTTCG ACACCCCTACC ACAATGCAAC
      ATCACTATGC TTAGATGTTG AGCTATGATT GATGTGGGAG TCGGGAAGC TGTGGGATGG TGTACGTTG

1261  GGTGTGTGAAG TACACGGTGG ATATTATATT GGATGGGTCT CCGTCCAGGA CCAAGTCGAG TCGCTTGTCA
      CCAACACTTC ATGTGCCACC TATAATATAA CCTACCCAGA GGCAGGTCCT GGTTCAGCTC AGCGAACAGT

1331  AACAGCAGGT TAGCCAGTAT CCGGACTACG CGCTGACCGT GACCGGCCAC KCCCTCGGCG CCTCCCTGGC
      TTGTCGTCCA ATCGGTCATA GGCTTGATGC GCGACTGGCA CTGGCCGGTG MGGGAGCCGC GGAGGGACCG

1401  GGCACCTCACT GCCGCCCAGC TGCTTGGCAG ATACGACAAC ATCCGCCCTGT ACACCTTCGG CGAACCCGCGC
      CCGTGAGTGA CGGCGGGTCG ACAGACGCTG TATGCTGTG TAGGCGGACA TGTGGAAGCC GCTTGGCGCG

      XhoI
      ~~~~~
1471  AGCGGCAATC AGGCCTTCGC GTCGTACATG AACGATGCCT TCCAAGCCTC GAGCCCAGAT ACGACGCAGT
      TCGCCGTTAG TCCGGAAGCG CAGCATGTAC TTGCTACGGA AGGTTCCGGAG CTCGGGTCTA TGCTGCCGTC

      NcoI
      ~~~~~
1541  ATTTCCGGGT CACTCATGCC AACGACGGCA TCCCAAACCT GCCCCCGGTG GAGCAGGGGT ACGCCCATGG
      TAAAGGCCCA GTGAGTACGG TTGCTGCCGT AGGTTTGA CGGGGGCCAC CTCGTCCCCA TCGGGGTACC

1611  CGGTGTAGAG TACTGGAGCG TTGATCCTTA CAGCGCCCAG AACACATTG TCTGCACCTGG GGATGAAGTG
      GCCACATCTC ATGACCTCGC AACTAGGAAT GTCGCGGGTC TTGTGTAAC AGACGTGACC CCTACTTCAC

1681  CAGTGCTGTG AGGCCCAGGG CGGACAGGGT GTGAATAATG CGCACACGAC TTATTTTGGG ATGACGAGCG
      GTCACGACAC TCCGGGTCCC GCCGTGTCCT CACTTATTAC GCGTGTCTG AATAAAACCC TACTGCTCGC

1751  GAGCCTGTAC ATGGTGATCA GTCATTTTCAG CCTCCCCGAG TGTACCAGGA AAGATGGATG TCCTGGAGAG
      CTCGGACATG TACCACCTAGT CAGTAAAGTC GGAGGGGCTC ACATGGTCTT TTCTACCTAC AGGACCTCTC

```

FIG._48D

```

1821 GGGGCCGCGT AACCACTGAA GGATGAGCTG TAAAGAAGCA GATCGTTCAA ACATTGGGCA ATAAAGTTTC
      CCCC GGCGCA TTGGTGACTT CCTACTCGAC ATTTCTTCGT CTAGCAAGTT TGTAACCCGT TATTCAAAG

1891 TTAAGATTGA ATCCTGTTGC CGGTCTTGGC ATGATTATCA TATAATTCTT GTTGAATTAC GTTAAGCATG
      AATTCTAACT TAGGACAAAC GCCAGAACGC TACTAATAGT ATATTAAAGA CAACTTAATG CAATTCGTAC

1961 TAATAATTAA CATGTAATGC ATGACGTTAT TTATGAGATG GGTTTTATG ATTAGAGTCC CGCAATTATA
      ATTATTAATT GTACATTACG TACTGCAATA AATACTCTAC CAAAATAAC TAATCTCAGG GCGTTAATAT

      BssHII          BssHII
      ~~~~~          ~~~~~

2031 CATTTAATAC GCGATAGAAA ACAAATATA GCGCGCAAAC TAGGATAAAT TATCGCGCGC GGTGTCATCT
      GTAAATTATG CGCTATCTTT TGTTTATAT CGCGCGTTTG ATCCTATTTA ATAGCGCGCG CCACAGTAGA

      XbaI
      ~~~~~

      ClaI HindIII
      ~~~~~

2101 ATGTTACTAG ATCGATAAGC TTCTAGAGCG GCCGGTGGAG CTCCAATTGC CCCTATAGTG AGTCGTATTA
      TACAATGATC TAGCTATTTC AAGATCTCGC CGGCCACCTC GAGGTTAAGC GGGATATCAC TCAGCATAAT

      BssHII
      ~~~~~

2171 CGCGCGCTCA CTGGCCGTCG TTTTACAACG TCGTGACTGG GAAACCCCTG GCGTTACCCA ACTTAATCGC
      GCGCGCGAGT GACCGGCAGC AAAATGTTGC AGCACTGACC CTTTGGGAC CGCAATGGGT TGAATTAGCG

2241 CTTGCAGCAC ATCCCCCTTT CGCCAGCTGG CGTAATAGCG AAGAGGCCCG CACCGATCGC CCTTCCCAAC
      GAACGTCGTG TAGGGGGAAA GCGGTCGACC GCATTATCGC TTCTCCGGGC GTGGCTAGCG GGAAGGGTTG

2311 AGTTGCGCAG CCTGAATGGC GAATGGGACG CGCCCTGTAG CGGCGCATT AAGCGCGCGG GTGTGGTGGT
      TCAACGCGTC GGACTTACCG CTTACCCTGC GCGGGACATC GCCGCGTAAT TCGCGCCGCC CACACCACCA

2381 TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT TCGCTTCTTT CCCTTCCTTT
      ATGCGCGTCG CACTGGCGAT GTGAACGGTC GCGGGATCGC GGGCGAGGAA AGCGAAAGAA GGAAGGAAA
  
```

FIG. 48E

142 / 154

2451 CTCGCCACGT TCGCCGGGCTT TCCCCGTCAA GCTCTAAATC GGGGGCTCCC TTTAGGGTTC CGATTTAGTG
GAGCGGTGCA AGCGGCCGAA AGGGCAGTT CGAGATTTAG CCCCCGAGG AATCCCAAG GCTAAATCAC

2521 CTTTACGGCA CCTCGACCCC AAAAAACTTG ATTAGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA
GAAATGCCGT GGAGCTGGGG TTTTITGAAC TAATCCCACT ACCAAGTGA TCACCCGGTA GCGGGACTAT

2591 GACGGTTTTC CGCCCTTTGA CGTTGGAGTC CACGTTCTTT AATAGTGGAC TCTTGTCCA AACTGGAACA
CTGCCAAAA GCGGAACT GCAACCTCAG GTGCAAGAAA TTATCACCTG AGAACCAAGGT TTGACCTTGT

2661 ACACCTCAACC CTATCTCGGT CTATCTCTTT GATTATAAG GGATTTTGCC GATTTCGGCC TATTGGTTAA
TGTGAGTTGG GATAGAGCCA GATAAGAAA CTAAATATTC CCTAAAAACGG CTAAAGCCGG ATAACCAATT

2731 AAAATGAGCT GATTAAACAA AAATTTAACG CGAATTTTAA CAAAATATTA ACGCTTACAA TTTAGGTGGC
TTTTTACTCGA CTAAATTGTT TTTAAATTGC GCTTAAATTT GTTTTATAAT TCGGAATGTT AAATCCACCG

2801 ACTTTTCGGG GAAATGTGCG CGGAACCCCT ATTTGTTTAT TTTTCTAAAT ACATTCAAAT ATGTATCCCG
TGAAAAAGCCC CTTTACACGC GCCTTGGGA TAAACAAATA AAAAGATTTA TGTAAGTTTA TACATAGGCG

2871 TCATGAGACA ATAACCCCTGA TAAATGCTTC AATAATATTG AAAAAGGAAG AGTATGAGTA TTCAACATTT
AGTACTCTGT TATTGGGACT ATTTACGAAG TTATTATAAC TTTTTCCTTC TCATACTCAT AAGTTGTAAA

2941 CCGTGTCCGC CTTATTCCCT TTTTTCGGC ATTTTGCCCTT CCTGTTTTG CTCACCCAGA AACGCTGGTG
GGCACAGCGG GAATAAGGGA AAAAACGCCG TAAAACGGAA GGACAAAAAC GAGTGGGTCT TTGCGACCAC

3011 AAAGTAAAG ATGCTGAAGA TCAGTTGGGT GCACGAGTGG GTTACATCGA ACTGGATCTC AACAGCGGTA
TTTCATTTTC TACGACTTCT AGTCAACCCA CGTGCTCACC CAATGTAGCT TGACCTAGAG TTGTCCGCCAT

3081 AGATCCTTGA GAGTTTTCGC CCCGAAGAAC GTTTTCCAAT GATGAGCACT TTTAAAGTTC TGCTATGTGG
TCTAGGAAC CTCAAAAAGC GGGCTTCTTG CAAAAGGTTA CTACTCGTGA AAATTTCAAG ACGATACACC

3151 CGCGGTATTA TCCCGTATTG ACGCCGGGCA AGAGCAACTC GGTCCGCCGA TACACTATTC TCAGAAATGAC
GCGCCATAAT AGGCGATAAC TGCGGCCCGT TCCTCGTTGAG CCAGCGCGCT ATGTGATAAG AGTCTTACTG

3221 TTGGTTGAGT ACTCACCAGT CACAGAAAAG CATCTTACGG ATGGCATGAC AGTAAAGAGAA TTATGCAGTG
AACCAACTCA TGAGTGGTCA GTGTCTTTTC GTAGAATGCC TACCGTACTG TCATTCTCTT AATACGTAC

FIG._48F

143 / 154

```

3291 CTGCCATAAC CATGAGTGAT AACACTGCGG CCAACTTACT TCTGACAACG ATCGGAGGAC CGAAGGAGCT
    GACGGTATTG GTACTCACTA TTGTGACGCC GGTGAATGA AGACTGTTGC TAGCCTCCTG GCTTCCTCGA

3361 AACCGCTTTT TTGCACAACA TGGGGGATCA TGTAACCTCG CTTGATCGTT GGAACCCGGA GCTGAATGAA
    TTGGCGAAAA AACGTGTTGT ACCCCCTAGT ACATTGAGCG GAACTAGCAA CCCTTGGCCT CGACTTACTT

3431 GCCATACCAA ACGACGAGCG TGACACCACG ATGCCGTGTAG CAATGGCAAC AACGTTGCGC AAACCTATTAA
    CGGTATGGTT TGCTGCTCGC ACTGTGGTGC TACGGACATC GTTACCCTTG TTGCAACGCG TTGATATAAT

3501 CTGGCGAACT ACTTACTCTA GCTTCCCGGC AACAAATTAAT AGACTGGATG GAGGCGGATA AAGTTGCAGG
    GACCGCTTGA TGAATGAGAT CGAAGGGCCG TTGTTAATTA TCTGACCTAC CTCCGCCCTAT TTCAACGTCC

3571 ACCACTTCTG CGCTCGGCCC TTCCGGCTGG CTGGTTTATT GCTGATAAAT CTGGAGCCGG TGAGCGTGGG
    TGGTGAAGAC GCGAGCCGGG AAGGCCGACC GACCAATAA CGACTATTA GACCTCGGCC ACTCGCACCC

3641 TCTCGCGGTA TCATTGCAGC ACTGGGGCCA GATGGTAAGC CCTCCCGTAT CGTAGTTATC TACACGACGG
    AGAGCGCCAT AGTAACGTCG TGACCCCGGT CTACCATTCTG GGAGGGCATA GCATCAATAG ATGTGCTGCC

3711 GGAGTCAGGC AACTATGGAT GAACGAAATA GACAGATCGC TGAGATAGGT GCCTCACTGA TTAAGCATTG
    CCTCAGTCCG TTGATACCTA CTTGCTTTAT CTGTCTAGCG ACTCTATCCA CGGAGTGACT AATTCGTAAC

3781 GTAACGTCA GACCAAGTTT ACTCATATAT ACTTTAGATT GATTAAAAA TTTCATTTTAA ATTTAAAAAGG
    CATTGACAGT CTGGTTCAAA TGAGTATATA TGAAATCTAA CTAAATTTTG AAGTAAAAAT TAAATTTTCC

3851 ATCTAGGTGA AGATCCTTTT TGATAAATCTC ATGACCAAAA TCCCTTAACG TGAGTTTTCG TTCCACTGAG
    TAGATCCACT TCTAGGAAAA ACTATTAGAG TACTGGTTT AGGGAATTGC ACTCAAAAGC AAGGTGACTC

3921 CGTCAGACCC CGTAGAAAAA ATCAAAAGGAT CTTCTTGAGA TCCTTTT TTTT CTGCGCGTAA TCTGCTGCTT
    GCAGTCTGGG GCATCTTTTC TAGTTTCCCTA GAAGAACTCT AGGAAAAAAA GACGCGCATT AGACGACGAA

3991 GCAAAACAAA AAACCAACCGC TACCAGCGGT GGTGTTGTTG CCGGATCAAG AGCTACCAAC TCTTTTTCGG
    CGTTTGTGTT TTTGGTGGCG ATGGTCGCCA CCAACCAAAAC GGCTTAGTTC TCGATGGTTG AGAAAAAGGC

4061 AAGGTAAC TGCTTCAGCAG AGCGCAGATA CCAAATACTG TCCTTCTAGT GTAGCCGTAG TTAGGCCACC
    TTCCATTGAC CGAAGTCGTC TCGCGTCTAT GGTATTATGAC AGGAAGATCA CATCGGCATC AATCCGGTGG

```

FIG. 48G

144 / 154

4131 ACTTCAAGAA CTCTGTAGCA CCGCCTACAT ACCTCGCTCT GCTAATCCTG TTACCAGTGG CTGCTGCCAG
TGAAGTTCTT GAGACATCGT GCGGATGTA TGGAGCGAGA CGATTAGGAC AATGGTCACC GACGACGGTC

4201 TGGCGATAAG TCGTGTCTTA CCGGGTTGGA CTCAAGACGA TAGTTACCGG ATAAGGCGCA GCGGTCGGGC
ACCGCTATT AGCACAGAAAT GGGCCAACTT GAGTTCTGCT ATCAATGGCC TATTCCGGCT CGCCAGCCCG

4271 TGAACGGGG GTTCGTGCAC ACAGCCCAGC TTGGAGCGAA CGACCTACAC CGAACTGAGA TACCTACAGC
ACTTGCCCC CAAGCACGTG TGTCGGGTG AACCTCGCTT GCTGGATGTG GCTTGACTCT ATGGATGTG

4341 GTGAGCTATG AGAAAGCGC ACGCTTCCC AGGGGAAAC GCCTGGTATC TTTATAGTCC TGTCGGGTTT
CACTCGATAC TCTTTCGGG TCGAAGGC TTCCCTCTTT CGGACCATAG AAATATCAGG ACAGCCCAA

4411 CGGAACAGGA GAGCGCACGA GGGAGCTTCC AGGGGAAAC GCCTGGTATC TTTATAGTCC TGTCGGGTTT
GCCTTGTCTT CTCGCGTGCT CCTCGAAG TCCCCCTTTG CGGACCATAG AAATATCAGG ACAGCCCAA

4481 CGCCACCTCT GACTTGAGCG TCGATTTTG TGATGCTCGT CAGGGGGCG GAGCCTATGG AAAAAAGCA
GCGGTGGAGA CTGAACTCGC AGCTAAAAAC ACTACGAGCA GTCCCCCGC CTCGGATACC TTTTTCGGT

4551 GCAACGCGG CTTTTTACGG TTCTTGGCCT TTTGCTGCC TTTTGCTCAC ATGTTCTTTC CTGCGTTATC
CGTTGCGCG GAAAAATGCC AAGGACCGA AAACGACCG AAACGAGTG TACAAGAAAG GACGCAATAG

4621 CCCTGATTCT GTGGATAACC GTATTACCG CTTTGAGTGA GCTGATACCG CTCGCCGCG CCGAACGACC
GGGACTAAGA CACCTATTGG CATAATGGC GAAACTCACT CGACTATGGC GAGCGCGCTC GGTTCGCTGG

4691 GAGCGCAGC AGTCAGTGAG CGAGGAAGC GAAGAGCGC CAATACGCA ACCGCCTCTC CCGCGCGGTT
CTCGCGTCG TCAGTCACTC GTCCTTCGC CTTCTCGCG GTTATGCGTT TGGCGGAGAG GGGCGCGCAA

4761 GGCCGATTCA TTAATGCAGC TGGCACGACA GGTTCCTCGA CTGGAAGCG GGCAGTGAGC GCAACGCAAT
CCGGCTAAGT AATTACGTG ACCGTGCTGT CCAAAGGGCT GACCTTTCCG CCGTCACTCG CGTTGCGTTA

4831 TAATGTGAGT TAGCTCACTC ATTAGGCACC CCAGGCTTTA CACTTTATGC TTCCGGCTCG TATGTTGTGT
ATTACACTCA ATCGAGTGAG TAATCCGTGG GGTCCGAAAT GTGAAATACG AAGGCCGAGC ATACAACACA

FIG._48H

4901 GGAATTGTGA GCGGATAACA ATTTACACACA GGAACACAGCT ATGACCATGA TTACGCCAAG CGCGCAATTA
CCTTAACACT CGCCTATTGT TAAAGTGTGT CCTTGTCTGA TACTGGTACT AATGCGGTTC GCGCGTTAAT

BssHII
~~~~~

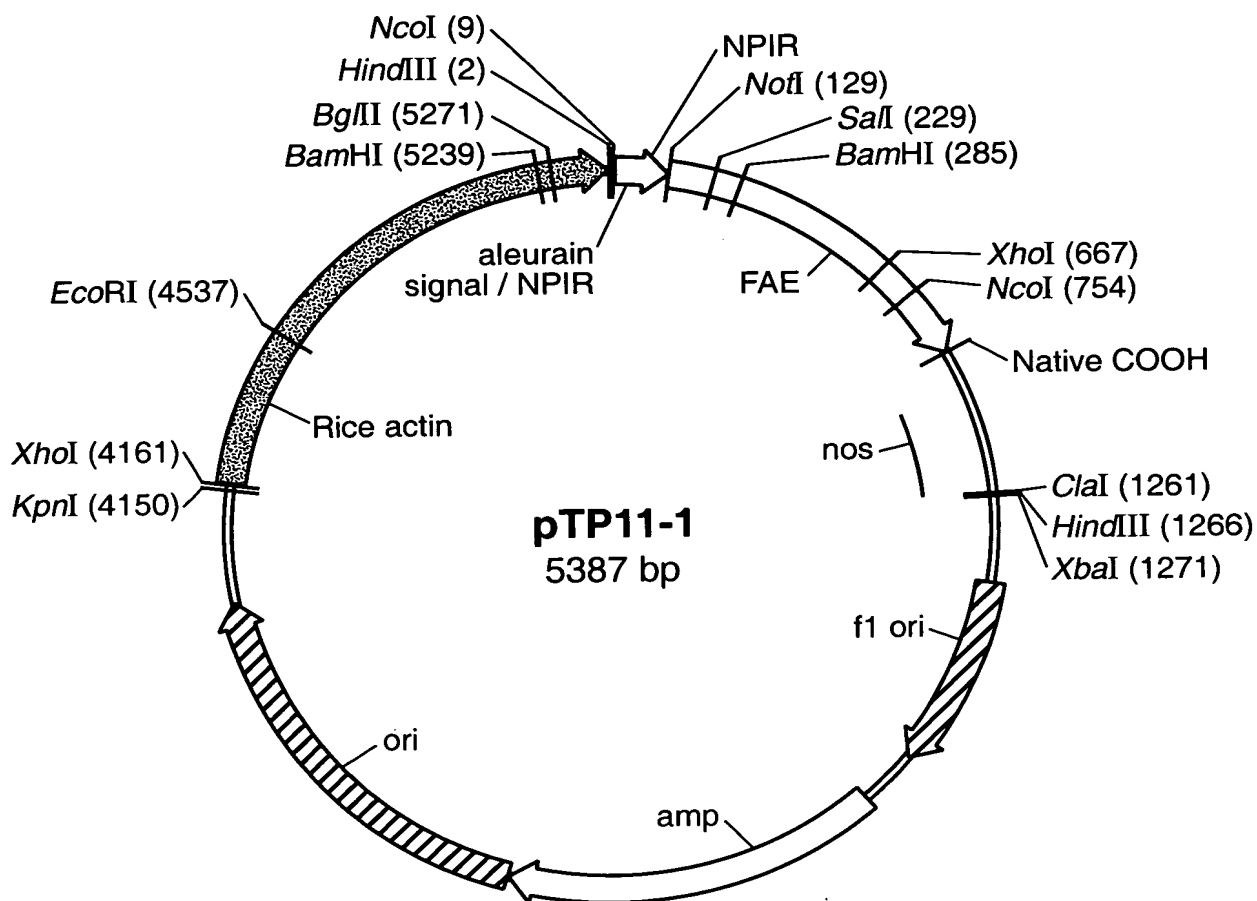
4971 ACCCTCACTA AAGGGAACAA AAGCTGGGTA C  
TGGGAGTGAT TTCCCTTGTT TTCGACCCAT G

NcoI  
~~~~~

KpnI
~~~~~

FIG.\_48I

146 / 154



**FIG.\_49A**

147 / 154

```

NcoI
~~~~~
HindIII
~~~~~
      M A H A R V L L L A L A V L A T A A V A V
1  AAGCTTACCA TGGCCACAGC CCGCGTCCTC CTCCTGGGCG TCGCCGTGCT GGCACGGCC GCCGTCGCCG
      NotI
      ~~~~~
 . A S S S F A D S N P I R P V T D R A A A S T .
71 TCGCCTCCTC CTCCTCCTTC GCCGACTCCA ACCCGATCCG GCCCGTCACC GACCGCGGG CCGCCTCCAC
 . Q G I S E D L Y S R L V E M A T I S Q A A Y A
141 GCAGGGCATC TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCACTA TCTCCAAGC TGCCTACGCC
 SalI
      ~~~~~
      ACCI
      ~~~~~
 D L C N I P S T I I K G E K I Y N S Q T D I N G
211 GACCTGTGCA ACATTCCGTC GACTATTATC AAGGAGAGA AAATTTACAA TTCTCAAACT GACATTAACG
 BamHI
      ~~~~~
      . W I L R D D S S K E I I T V F R G T G S D T N .
281 GATGGATCCT CCGCGACGAC AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCACCTGGTA GTGATACGAA
      . L Q L D T N Y T L T P F D T L P Q C N G C E V
351 TCTACAATC GATACTAAT ACACCTTCAC GCCTTTCGAC ACCCTACCAC AATGCAACGG TTGTGAAGTA
      H G G Y Y I G W V S V Q D Q V E S L V K Q Q V S
421 CACGGTGGAT ATTATATTGG ATGGGTCTCC GTCCAGGACC AAGTCGAGTC GCTTGTCAAA CAGCAGGTTA
      . Q Y P D Y A L T V T G H X L G A S L A A L T A .
491 GCCAGTATCC GGACTACGG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG CACTCACTGC
      . A Q L S A T Y D N I R L Y T F G E P R S G N Q
561 CGCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC ACCTTCGGCG AACCGGCGAG CGGCAATCAG
      XhoI
      ~~~~~
 A F A S Y M N D A F Q A S S P D T T Q Y F R V T
631 GCCTTCGGGT CGTACATGAA CGATGCCTTC CAAGCCTCGA GCCCAGATAC GACGCAGTAT TTCCGGGTCA

```

FIG. 49B

148 / 154

```

 . H A N D G I P N L P P V E Q G Y A H G G V E Y .
701 CTCATGCCAA CGACGGCATC CCAAACCTGC CCCCGGTGGA GCAGGGGTAC GCCCATGGCG GTGTAGAGTA
 . W S V D P Y S A Q N T F V C T G D E V Q C C E
771 CTGGAGCGTT GATCCTTACA GCGCCAGAA CACATTGTC TGCACCTGGG ATGAAGTGCA GTGCTGTGAG
 A Q G G Q G V N N A H T T Y F G M T S G A C T W
841 GCCCAGGCG GACAGGGTGT GAATAATGCG CACACGACTT ATTTGGGAT GACGAGCGGA GCCTGTACAT
 . *
911 GGTGATCAGT CATTTAGCC TCCCCGAGTG TACCAGGAAA GATGGATGTC CTGGAGAGGG GGCCGCGTAA
981 CCACTGAAGG ATGAGCTGTA AAGAAGCAGA TCCTTCAAAC ATTTGGCAAT AAAGTTCTT AGGATTGAAT
1051 CCTGTGCGG GTCTTGCGAT GATTATCATA TAATTCTGT TGAATTACGT TAAGCATGTA ATAATFAACA
1121 TGTAAATGCAT GACGTTATTT ATGAGATGGG TTTTATATGAT TAGAGTCCCG CAATTATACA TTTAATACGC
 ClaI
1191 GATAGAAAAC AAAATATAGC GCGCAAACTA GGATAAATTA TCGCGCGCGG TGTATCTAT GTTACTAGAT
 HindIII
 ClaI XbaI
      ~~~~~
1261 CGATAAGCTT CTAGAGCGGC CGGTGGAGCT CCAATTCGCC CTATAGTGAG TCGTATTACG CGCGCTCACT
1331 GGCCGTGCTT TTACAACGTC GTGACTGGGA AACCCCTGGC GTTACCCAAC TTAATCGCCT TGACGACACAT
1401 CCCCTTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC TTCCCAACAG TTGCGCAGCC
1471 TGAATGGCGA ATGGGACGCG CCTGTAGCG CGCTCCTTTC GCTTCTTTCG GTGGTGGTTA CGCGCAGCGT
1541 GACCGGTACA CTTGCCAGCG CCTAGCGCC GGGCTCCCTT TAGGGTTCCG ATTTAGTGCT TTACGGCACC
1611 GCGGCTTTC CCCGTCAAAG TCTAAATCGG GTTCAACGTC TAGGGCCATCG CCCTGATAGA CCGTTTTCG
1681 TCGACCCCAA AAACTTGAT TAGGTGATG CGTCTTTAA TAGTGACATC TTGTTCCAAA CTGGAACAACT ACTCAACCTT
1751 CCCTTTGACG TTGGAGTCCA ATCTTTTGA TTTATAAGGG ATTTTGCCGA TTTCGGCCTA TTGGTTAAAA AATGAGCTGA
1821 ATCTCGGTCT ATCTTTTGA TTTATAAGGG ATTTTGCCGA TTTCGGCCTA TTGGTTAAAA AATGAGCTGA
1891 TTTAACAAAA ATTTAACGCG AATTTTAACTA AATATTAAC GCTTAAATAT ATTCAATAT GTATCCGCTC ATGAGACAAAT
1961 AATGTGCGG GAACCCCTAT TTGTTTATTT TTCTAAATAC ATTCAATAT TATGAGTATT CAACATTTCC GTGTCGCCCT
2031 AACCCTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT CAACATTTCC GTGTCGCCCT
2101 TATTCCTTTT TTTGCGGCAT TTTGCTTCC TGTTTTGTCT CACCCAGAAA CGCTGGTGAA AGTAAAAAGAT
2171 GCTGAAGATC AGTTGGGTGC ACAGTGGGT TACATCGAAC TGGATCTCAA CAGCGGTAA GATCCTTGAGA
2241 GTTTTCGCC CCAAGAACGT TTTCCAATGA TGAGCACTTT TAAAGTTCTG TATGTTGGCG CCGTATTATC
2311 CCGTATTGAC GCCGGGCAAG AGCAACTCGG TCGCCGCATA CACTATTCTC AGAATGACTT GGTGAGTAC
2381 TCACCAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG TAAGAGAATT ATGCAGTGCT GCCATAACCA

```

FIG. 49C

149 / 154

```

2451  TGAGTGATAA  CACTGCGGCC  AACTTACTTC  TGACAAACGAT  CGGAGGACCG  AAGGAGCTAA  CCGCTTTTTC  CCGCTTTTTC
2521  GCACAACATG  GGGGATCATG  TAACTCGCCT  TGATCGTTGG  GAACCGGAGC  TGAATGAAGC  CATACCAAAC  CATACCAAAC
2591  GACGAGCGTG  ACACCACGAT  GCGTGTAGCA  ATGGCAACAA  CGTTGCGCAA  ACTATTAACT  GGCGAACTAC  GGCGAACTAC
2661  TTACTCTAGC  TTCCCGGCAA  CAATTAATAG  ACTGGAATGA  GCGGGATAAA  GTTGCAGGAC  CACTTCTGCG  CACTTCTGCG
2731  CTCGGCCCTT  CCGGCTGGCT  GGTTTATTGC  TGATAAATCT  GGAGCCGGTG  AGCGTGGGTC  TCGCGGTATC  TCGCGGTATC
2801  ATTGCAGCAC  TGGGGCCAGA  TGGTAAGCCC  TCCCGTATCG  TAGTTATCTA  CACGACGGGG  AGTCAGGCAA  AGTCAGGCAA
2871  CTATGGATGA  ACGAAATAGA  CAGATCGCTG  AGATAGGTGC  CTCACGTGAT  AAGCATTTGT  AACTGTCAGA  AACTGTCAGA
2941  CCAAGTTTAC  TCATATATAC  TTTAGATTGA  TTTAAAACTT  CATTTTAAAT  TTAAAAAGGAT  CTAGGTGAAG  CTAGGTGAAG
3011  ATCCTTTTTC  ATAATCTCAT  GAACAAAATC  CCTTAAACGT  AGTTTTCGTT  CCACGTAGCG  TCAGACCCCG  TCAGACCCCG
3081  TAGAAAAGAT  CAAAGGATCT  TCTTGAGATC  CTTTTTTCT  GCGCGTAATC  TGCTGCTTGC  AAACAAAATA  AAACAAAATA
3151  ACCACCGCTA  CCAGCGGTGG  TTTGTTTGCC  GGATCAAGAG  CTACCAACTC  TTTTTCGAA  GGTAACCTGG  GGTAACCTGG
3221  TTCAGCAGAG  CGCAGATACC  AAATACTGTC  CTTCTAGTGT  AGCCGTAGTT  AGGCCACCAC  TTCAAGAACT  TTCAAGAACT
3291  CTGTAGCACC  GCCTACATAC  CTCGCTCTGC  TAATCCTGTT  ACCAGTGGCT  GCTGCCAGTG  GCGATAAGTC  GCGATAAGTC
3361  GTGCTTACC  GGGTTGGACT  CAAGACGATA  GTTACCGGAT  AAGGCGCAGC  GGTGCGGCTG  AACGGGGGT  AACGGGGGT
3431  TCGTGCACAC  AGCCAGCTT  GGAGCGAAG  ACCTACACCG  AACTGAGATA  CCTACAGCGT  GAGCTATGAG  GAGCTATGAG
3501  AAAGCGCCAC  GCTTCCCGAA  GGGAGAAAG  CGGACAGGTA  TCCGGTAAGC  GGCAGGTCG  GAACAGGAGA  GAACAGGAGA
3571  GCGCAGCAGG  GAGCTTCCAG  GGGGAAACGC  CTGGTATCTT  TATAGTCTTG  TCGGGTTTCG  CCACCTCTGA  CCACCTCTGA
3641  CTTGAGCGTC  GATTTTGTG  ATGCTCGTCA  GGGGGCGGA  GCCTATGGAA  AAACGCCAGC  AACGCGGCTT  AACGCGGCTT
3711  TTTTACGGTT  CCTGGCCCTT  TGCTGGCCTT  TTGCTCACAT  GTTCTTTTCT  GCGTTATCCC  CTGATTCTGT  CTGATTCTGT
3781  GGATAACCGT  ATTACCGCTT  TTGAGTGAGC  TGATACCGCT  CGCCGACGCC  GAACGACCGA  CGCAGCGGAG  CGCAGCGGAG
3851  TCAGTGAGCG  AGGAAGCGGA  AGAGCGCCCA  ATACGCAAA  CGCCTCTCCC  CGCGGTTGG  CCGATTCAAT  CCGATTCAAT
3921  AATGCAGCTG  GCACGACAGG  TTTCCCGACT  GGAAGCGGG  CAGTGAGCGC  AACGCAATTA  ATGTGAGTTA  ATGTGAGTTA
3991  GCTCACTCAT  TAGGCACCCC  AGGCTTTACA  CTTTATGCTT  CCGGCTCGTA  TGTTGTGTGG  AATTGTGAGC  AATTGTGAGC
4061  GGATAACAAT  TTCACACAGG  AACACAGCTA  GACCATGATT  ACGCCAAGCG  CGCAATTAAC  CCTCACTAAA  CCTCACTAAA

                                KpnI
~~~~~
4131 GGGAACAAA GCTGGGTACC GGGCCCCCCC TCGAGGTCAT TCATATGCTT GAGAAAGAGG TCGGGATAGT TCGGGATAGT
4201 CCAAAATAA ACAAAAGTAA GATTACCTGG TCAAAAGTGA AAACATCAGT TAAAAAGTGG TATAAGTAAA TATAAGTAAA
4271 ATATCGGTAA TAAAAAGTGG CCAAAAGTGA AATTACTCT TTTCTACTAT TATAAAAAAT GAGGATGTTT GAGGATGTTT
4341 TGTCGGTACT TTGATACGTC ATTTTGTAT GAATTGGTT TTAAGTTTAT TCGCGATTGG GAAATGCATA GAAATGCATA
4411 TCTGTATTG AGTCGGTTT TAAGTTCGTT GCTTTGTAA ATACAGAGGG ATTTGTATAA GAAATATCTT GAAATATCTT

 XhoI
~~~~~
4481  TAAAAAACCC  ATATGCTAAT  TTGACATAAT  TTTTGAGAAA  AATATATATT  CAGGCGAATT  CCACAATGAA  CCACAATGAA
4551  CAATAATAAG  ATTAAAAATG  CTTGCCCCCC  TTGCAGCGAT  GGGTATTTTT  TCTAGTAAAA  TAAAAAGATA  TAAAAAGATA
4621  ACTTAGACTC  AAAACATTTA  CAAAAACAAC  CCTAAAGTC  CTAAAGCCCC  AAGTGCCTATG  CACGATCCAT  CACGATCCAT

```

FIG. 49D

```

4691 AGCAAGCCCA GCCCAACCCA ACCCAACCCA ACCCAACCCA GTGCAGCCAA CTGGCAATA GTCTCCACCC
4761 CCGGCACTAT CACCGTGAGT TGTCCGCACC ACGCACGTC TCGCAGCCAA AAAAAAAA AGAAGAAAA
4831 AAAAGAAAA GAAAAACAGC AGGTGGGTCC TCGTGTGGG GCGCGGAAA GCGAGGAGGA TCGCGAGCAG
4901 CGACGAGGCC CGGCCCTCCC TCCGCTTCCA AAGAAACGCC CCCCATCGCC ACTATATACA TACCCCCCCC
4971 TCTCCTCCCA TCCCCCACA CCTACCACA CCACCAACAC CACCTCCTCC CCCCTCGCTG CCGGACGACG
5041 AGTCCTCCC CCTCCCCCT CCGCCGCCGC CGGTAAACAC CCGCCCCCTC TCCTCTTTCT TTCTCCGTTT
5111 TTTTTTTCGT CTCGGTCTCG ATCTTTGGCC TTGGTAGTTT GGGTGGGCGA GAGCGGCTTC GTCGCCCAGA

                                     BamHI
                                     ~~~~~

5181 TCGGTGCGCG GGAGGGGCGG GATCTCGCGG CTGGCGTCTC CGGGCGTGAG TCGGCCCGGA TCCTCGCGGG

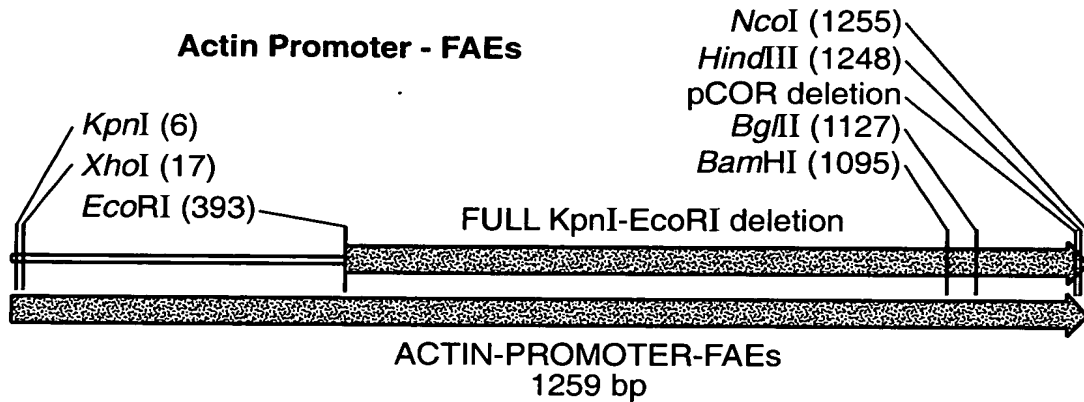
 BglII
                                     ~~~~~

5251 GAATGGGGCT CTCGGATGTA GATCTTCTTT CTTTCTTCTT TTTGTGGTAG AATTGAATC CCTCAGCAT
5321 GTTCATCGGT AGTTTTCCTT TTCATGATTT GTGACAAATG CAGCCTCGTG CGGAGCTTTT TTGTAGC

```

FIG.\_49E

151 / 154



|     | KpnI               | XhoI              |                   |                    |                   |                   |                   |
|-----|--------------------|-------------------|-------------------|--------------------|-------------------|-------------------|-------------------|
|     | ~~~~~              | ~~~~~             |                   |                    |                   |                   |                   |
| 1   | <u>GGTACCGGGC</u>  | <u>CCCCCTCGA</u>  | <u>GGTCATTCAT</u> | <u>ATGCTTGAGA</u>  | <u>AGAGAGTCGG</u> | <u>GATAGTCCAA</u> | <u>AATAAAACAA</u> |
|     | <u>CCATGGCCCCG</u> | <u>GGGGGGAGCT</u> | <u>CCAGTAAGTA</u> | <u>TACGAACCTCT</u> | <u>TCTCTCAGCC</u> | <u>CTATCAGGTT</u> | <u>TTATTTTGTT</u> |
| 71  | <u>AGGTAAGATT</u>  | <u>ACCTGGTCAA</u> | <u>AAGTGAAAAC</u> | <u>ATCAGTTAAA</u>  | <u>AGGTGGTATA</u> | <u>AGTAAAATAT</u> | <u>CGGTAATAAA</u> |
|     | <u>TCCATTCTAA</u>  | <u>TGGACCAGTT</u> | <u>TTCACTTTTG</u> | <u>TAGTCAATTT</u>  | <u>TCCACCATAT</u> | <u>TCATTTTATA</u> | <u>GCCATTATTT</u> |
| 141 | <u>AGGTGGCCCA</u>  | <u>AAGTGAAATT</u> | <u>TACTCTTTTC</u> | <u>TACTATTATA</u>  | <u>AAAATTGAGG</u> | <u>ATGTTTTGTC</u> | <u>GGTACTTTGA</u> |
|     | <u>TCCACCGGGT</u>  | <u>TTCACTTTAA</u> | <u>ATGAGAAAAG</u> | <u>ATGATAATAT</u>  | <u>TTTAACTCC</u>  | <u>TACAAAACAG</u> | <u>CCATGAAACT</u> |
| 211 | <u>TACGTCATTT</u>  | <u>TTGTATGAAT</u> | <u>TGGTTTTTAA</u> | <u>GTTTATTCGC</u>  | <u>GATTTGGAAA</u> | <u>TGCATATCTG</u> | <u>TATTTGAGTC</u> |
|     | <u>ATGCAGTAAA</u>  | <u>AACATACTTA</u> | <u>ACCAAAAATT</u> | <u>CAAATAAGCG</u>  | <u>CTAAACCTTT</u> | <u>ACGTATAGAC</u> | <u>ATAAACTCAG</u> |
| 281 | <u>GGTTTTTAAG</u>  | <u>TTCGTTGCTT</u> | <u>TTGTAAATAC</u> | <u>AGAGGGATTT</u>  | <u>GTATAAGAAA</u> | <u>TATCTTTAAA</u> | <u>AAACCCATAT</u> |
|     | <u>CCAAAAATTC</u>  | <u>AAGCAACGAA</u> | <u>AACATTTATG</u> | <u>TCTCCCTAAA</u>  | <u>CATATTCTTT</u> | <u>ATAGAAATTT</u> | <u>TTTGGGTATA</u> |
|     |                    |                   |                   | <u>EcoRI</u>       |                   |                   |                   |
|     |                    |                   |                   | ~~~~~              |                   |                   |                   |
| 351 | <u>GCTAATTTGA</u>  | <u>CATAATTTTT</u> | <u>GAGAAAAATA</u> | <u>TATATTCAGG</u>  | <u>CGAATTCCAC</u> | <u>AATGAACAAT</u> | <u>AATAAGATTA</u> |
|     | <u>CGATTAAACT</u>  | <u>GTATTAAAAA</u> | <u>CTCTTTTTAT</u> | <u>ATATAAGTCC</u>  | <u>GCTTAAGGTG</u> | <u>TTACTTGTTA</u> | <u>TTATTCTAAT</u> |
| 421 | <u>AAATAGCTTG</u>  | <u>CCCCCGTTGC</u> | <u>AGCGATGGGT</u> | <u>ATTTTTTCTA</u>  | <u>GTAAAATAAA</u> | <u>AGATAAACTT</u> | <u>AGACTCAAAA</u> |
|     | <u>TTTATCGAAC</u>  | <u>GGGGGCAACG</u> | <u>TCGCTACCCA</u> | <u>TAAAAAAGAT</u>  | <u>CATTTTATTT</u> | <u>TCTATTTGAA</u> | <u>TCTGAGTTTT</u> |
| 491 | <u>CATTTACAAA</u>  | <u>AACAACCCCT</u> | <u>AAAGTCCTAA</u> | <u>AGCCCAAAGT</u>  | <u>GCTATGCACG</u> | <u>ATCCATAGCA</u> | <u>AGCCAGCCCC</u> |
|     | <u>GTAAATGTTT</u>  | <u>TTGTTGGGGA</u> | <u>TTTCAGGATT</u> | <u>TCGGGTTTCA</u>  | <u>CGATACGTGC</u> | <u>TAGGTATCGT</u> | <u>TCGGGTCGGG</u> |
| 561 | <u>AACCCAACCC</u>  | <u>AACCCAACCC</u> | <u>ACCCAGTGC</u>  | <u>AGCCAACTGG</u>  | <u>CAAATAGTCT</u> | <u>CCACCCCCGG</u> | <u>CACTATCACC</u> |
|     | <u>TTGGGTGGG</u>   | <u>TTGGGTGGG</u>  | <u>TGGGGTCACG</u> | <u>TCGGTTGACC</u>  | <u>GTTTATCAGA</u> | <u>GGTGGGGGCC</u> | <u>GTGATAGTGG</u> |
| 631 | <u>GTGAGTTGTC</u>  | <u>CGCACCACCG</u> | <u>CACGTCTCGC</u> | <u>AGCCAAAAAA</u>  | <u>AAAAAAAGAA</u> | <u>AGAAAAAA</u>   | <u>GAAAAAGAAA</u> |
|     | <u>CACTCAACAG</u>  | <u>GCGTGGTGGC</u> | <u>GTGCAGAGCG</u> | <u>TCGGTTTTTT</u>  | <u>TTTTTTTCTT</u> | <u>TCTTTTTTTT</u> | <u>CTTTTTCTTT</u> |
| 701 | <u>AACAGCAGGT</u>  | <u>GGGTCCGGGT</u> | <u>CGTGGGGGCC</u> | <u>GGAAAAGCGA</u>  | <u>GGAGGATCGC</u> | <u>GAGCAGCGAC</u> | <u>GAGGCCCGGC</u> |
|     | <u>TTGTCGTCCA</u>  | <u>CCCAGGCCCA</u> | <u>GCACCCCCCG</u> | <u>CCTTTTCGCT</u>  | <u>CCTCCTAGCG</u> | <u>CTCGTCGCTG</u> | <u>CTCCGGGCCG</u> |

**FIG. 50A**

152 / 154

771 CCTCCCTCCG CTTCCAAAGA AACGCCCCC ATCGCCACTA TATACATACC CCCCCTCTC CTCCCATCCC  
GGAGGGAGGC GAAGGTTTCT TTGCGGGGG TAGCGGTGAT ATATGTATGG GGGGGAGAG GAGGGTAGGG

841 CCCAACCCTA CCACCACCAC CACCACCACC TCCTCCCCC TCGTGCCGG ACGACGAGCT CCTCCCCCT  
GGGTGGGAT GGTGGTGGTG GTGGTGGTG AGGAGGGGG AGCGACGGCC TGCTGCTCGA GGAGGGGGGA

911 CCCCCTCCGC CGCCGCCGGT AACCACCCCG CCCCTCTCCT CTTTCTTTCT CCGTTTTTTT TTTCGTCTCG  
GGGGGAGGCG GCGGCGGCCA TTGGTGGGGC GGGGAGAGGA GAAAGAAAGA GGCAAAAAA AAAGCAGAGC

981 GTCTCGATCT TTGGCCTTGG TAGTTTGGGT GGGCGAGAGC GGCTTCGTCG CCCAGATCGG TCGCGGGAG  
CAGAGCTAGA AACCGGAACC ATCAAACCCA CCCGCTCTCG CCGAAGCAGC GGGTCTAGCC ACGCGCCCTC

BamHI

~~~~~

1051 GGGCGGGATC TCGCGGCTGG CGTCTCCGGG CGTGAGTCGG CCCGATCCT CGCGGGGAAT GGGGCTCTCG
CCCGCCCTAG AGCGCCGACC GCAGAGGCC GCACTCAGCC GGGCCTAGGA GCGCCCCTTA CCCCAGAGAGC

BglII

~~~~~

1121 GATGTAGATC TTCTTTCTTT CTTCTTTTGG TGGTAGAATT TGAATCCCTC AGCATTGTTC ATCGGTAGTT  
CTACATCTAG AAGAAAGAAA GAAGAAAAAC ACCATCTTAA ACTTAGGGAG TCGTAACAAG TAGCCATCAA

HindIII NcoI

~~~~~

1191 TTTCTTTTCA TGATTGTGA CAAATGCAGC CTCGTGCGGA GCTTTTTTGT AG**GTAGA**AAGC TTACCATGG
AAAGAAAAGT ACTAAACACT GTTTACGTCG GAGCACGCCT CGAAAAAACA TC**CATC**TTTCG AATGGTACC

KpnI-EcoRI - deletion underlined and restored NCO site in bold in vectors pJQ4.9,
pJQ3.2 and pJO6.3.

FIG._50B

153 / 154

ALEURAIN_d I t d NPIR (Ap plast) Structur and Sequ nc



ALEURAIN-NPIR-DEL
93 bp

+1 M A H A R V L L L A L A V L A T A A V A
HindIII NcoI

1 AAGCTTACCA TGGCCACGC CCGCGTCCTC CTCCTGGCGC TCGCCGTGCT GGCCACGGCC GCCGTGCGCG
TTCGAATGGT ACCGGGTGCG GGCGCAGGAG GAGGACCGCG AGCGGCACGA CCGGTGCCGG CGGCAGCGGC

+1 V A S S R A A
NotI

71 TCGCCTCCTC CCGCGCGGCC GCC
AGCGGAGGAG GGCGCGCCGG CGG

FIG._51

154 / 154

SEE1 (Senescence enhanced) PROMOTER sequence

```

1   CATGGGCCAG GTATAATTAT GGGATATCTC AAGCAAATAA TCGAAATATC ACCATTGGCT ACAATATCTG
      PstI              XbaI      XbaI
      ~~~~~
71  AGCTCCGAGT TCTGACTGCA GTCTGGATGA CGCGTGTGTG ATCTAGAACT CTAGATAGCA CAGCCACAGC
141 ACCTACAGGA GTGCGACACT TGTGGACTGT AGTAGTGTTG GAGACGGAGC TCTTTCCTAC CTCCTGACGT
211 TGCCGCCGTT GTCCATTCCA ACGGCATCAC TCTCAACCAA TCACGCGCTC CCAACAAAAT ATCGTCCCCC
281 ATGTCTTGGC GGAGAGAGAG TACATACATG CTGTCGCGCC GTTTTGTCT GAATCTCGCT TCCACTGGCC
      SmaI
      ~~~~~
351 AATCAGCTCA GCTCCCGGGA GCTCACTCAT TCAAGATCCC ATCGTCGTCG TCACCCCTGG CGTCATGGGA
421 TGGAAAAGAA CCTCCGTTGC TCGGATGAGT CAGCCATATC CCCGAACAGA GTACTGCAAG ATAACCCAAT
      SphI
      ~~~~~
491 TCAGATTCCC CCAATAGAGA AAGTATAGCA TGCTTTCGGG TTTTGTGTTG CTTAATTGAC TTTATTTTGT
561 TTGGAGTTGA ATGCTGATTT GTTGTGTAAA ATGCCCAACC ATCTGAATAT CGAGACGGAT AATAGGCTGG
631 CTAATTAATT TATAGCAAGA TTCTGTAGTG CACATCGCAA ATATCTTTCT GGGCATTACA GCTGGAGGCT
      PstI
      ~~~~~
701 TCATCAGCCT GAAACACTCT GCAGAGCCTG AAGCAAGTGG TGAAGCGTGG CGATGAGATG GGTATAAAAC
771 CCCCggcacc GGGACGCGAG CTCCCGCCTA CCAGTACCAT CTCGCCTCGC TCCCCCTGCC GGACGACCCA
841 GTAAAATACT GTTGCCCACT CGCCGGCGAG ATG

```

FIG._52

SEE1 (Senescence enhanced) PROMOTER plus vacuolar aleurain SIGNAL/NPIR sequence

```

1   CATGGGCCAG GTATAATTAT GGGATATCTC AAGCAAATAA TCGAAATATC ACCATTGGCT ACAATATCTG
      PstI              XbaI      XbaI
      ~~~~~
71  AGCTCCGAGT TCTGACTGCA GTCTGGATGA CGCGTGTGTG ATCTAGAACT CTAGATAGCA CAGCCACAGC
141 ACCTACAGGA GTGCGACACT TGTGGACTGT AGTAGTGTTG GAGACGGAGC TCTTTCCTAC CTCCTGACGT
211 TGCCGCCGTT GTCCATTCCA ACGGCATCAC TCTCAACCAA TCACGCGCTC CCAACAAAAT ATCGTCCCCC
281 ATGTCTTGGC GGAGAGAGAG TACATACATG CTGTCGCGCC GTTTTGTCT GAATCTCGCT TCCACTGGCC
      SmaI
      ~~~~~
351 AATCAGCTCA GCTCCCGGGA GCTCACTCAT TCAAGATCCC ATCGTCGTCG TCACCCCTGG CGTCATGGGA
421 TGGAAAAGAA CCTCCGTTGC TCGGATGAGT CAGCCATATC CCCGAACAGA GTACTGCAAG ATAACCCAAT
      SphI
      ~~~~~
491 TCAGATTCCC CCAATAGAGA AAGTATAGCA TGCTTTCGGG TTTTGTGTTG CTTAATTGAC TTTATTTTGT
561 TTGGAGTTGA ATGCTGATTT GTTGTGTAAA ATGCCCAACC ATCTGAATAT CGAGACGGAT AATAGGCTGG
631 CTAATTAATT TATAGCAAGA TTCTGTAGTG CACATCGCAA ATATCTTTCT GGGCATTACA GCTGGAGGCT
      PstI
      ~~~~~
701 TCATCAGCCT GAAACACTCT GCAGAGCCTG AAGCAAGTGG TGAAGCGTGG CGATGAGATG GGTATAAAAC
771 CCCCggcacc GGGACGCGAG CTCCCGCCTA CCAGTACCAT CTCGCCTCGC TCCCCCTGCC GGACGACCCA
      M A H G R I L F L A L A V L
841 GTAAAATACT GTTGCCCACT CGCCGGCGAG ATGGCCACG GCCGCATCCT CTTCTTGGCG CTCGCCGTCT
      BssHII
      ~~~~~
      NotI
      ~~~~~
      A T A A V A A A S L A D S N P I R P V T E R A
911 TGGCCACCGC CGCGGTGGCC GCCGCATCNT TGGCGGACTC CAACCCGATC CGGCCCGTCA CCGAGCGCGC
      NotI
      ~~~~~
      A A
981 GGCCGCC

```

FIG._53